Neural network modelling of hot deformation of austenite

Mathew Peet

Wolfson College

University of Cambridge

A thesis submitted for the degree of Master of Philosophy in Materials Modelling at the University of Cambridge.

August 2001

Index

Preface

This dissertation is submitted for the degree of Master of Philosophy in materials modelling at the University of Cambridge. The research described herein was conducted under the supervision of Prof. H. K. D. H. Bhadeshia in the Department of Materials Science and Metallurgy, Cambridge, between June 2001 and September 2001. This dissertation contains less than 15, 000 words.

Except where acknowledgement and reference is made to previous work, this work is, to the best of my knowledge, original. Neither this, nor any substantially similar dissertation has been, or is being, submitted for any other degree, diploma or other qualification at any other university.

Mathew Peet

Acknowledgements

I would like to thank Professor A. H. Windle for the provision of laboratory facilities in the department of Materials Science and Metallurgy at the University of Cambridge.

I would also like to express thanks to my supervisor Prof. H. K. D. H. Bhadeshia for his encouragement during this work. Also all those who have contributed to the M. Phil. course, either in organising the course in this first year or by lecturing.

Abstract

It is known that the hot deformation behaviour of austenite in steels is a complicated process, dependent on chemical composition, microstructure, temperature and strain rate. While many models have been developed to represent the flow stress as a function of these variables, it is however not yet possible to predict the behaviour for a new alloy. The effects of the different variables on the flow stress are investigated. Linear regression techniques are not capable of representing the data, however neural networks are capable of modelling highly non−linear data. A neural network model has been developed using a large database of compositions and meaningful inputs, including composition. The model allows the calculation of error bars that depend upon the position of a prediction in the input space and the level of perceived noise in the data. The validity of the model was evaluated by testing against six compositions of carbon−manganese steels.

Nomenclature

R	Gas Constant
k	Boltzman Constant
b	Burgers vector
exp	base of natural logarithms
exp_{10}	base 10 exponential function
Q	activation energy
ΔG	Free energy
ΔF	Helmholtz energy
F	force
A	area
log	logarithm to base 10
ln	natural logarithm
T	Temperature
T_m	Melting temperature
n	Strain rate sensitivity
σ	True Stress
ε	true strain
$\dot{\varepsilon}$	true−strain rate
a	lattice parameter, or physical constant
c	concentration
τ	Shear stress
$M_{(w)}$	Objective function in neural networks
$E_{(D)}$	Error function in neural networks
$E_{(w)}$	Regularisation term in neural network
σ_w	Model perceived significance
σ_{nu}	Model perceived noise in the output
α	Hyperparameter in neural networks, physical constant
β	Hyperparameter in neural networks
σ_0	Yield stress, or yield strength,
σ_0, ε^*, ε_0, n^*	phenomological material parameter in generalised power law
σ_s, σ_i, ?, ε_i, r	phenomological material parameter in saturation type equation
SFE	Stacking fault energy

Please note the atomic diameters in Table 2 are in Ångstoms at a temperature of 1000°C

1. Introduction

Overview

The vast majority of steels contain low concentrations of alloy elements, and are shaped using hot−rolling. Steels become softer with increasing temperature, so the working temperatures are within the austenitic phase field, the upper limit governed by practical limitations associated with reheating [1]. Models of hot working of steel rely on mathematical representations of the flow stress versus the plastic strain behaviour of austenite, including the effect of strain rate and temperature, these are known as constitutive equations.

Materials science has helped in understanding the phenomena that affect processing, but there remain many problems were quantitative treatments are dismally lacking [2]. Whereas theoretical predictions can sometimes be made of simple properties such as the yield strength of a microstructure using dislocation theory and others, it is not yet possible to predict the strain−hardening coefficient [2]. The lack of progress in predicting the mechanical properties is because of their dependence on a large number of variables. And yet, there are neural networks which are extremely useful in these circumstances, allowing a quantitative expression of mechanical properties for complex problems. Neural networks can successfully model complex problems in materials science, which seem overwhelming from a fundamental perspective and where simplification is unacceptable.

There are many examples of neural networks applied to modelling the processes of hot−rolling, demonstrating the applicability of this approach. Singh *et al*. estimated the yield strength and tensile strength of steel as a function of 108 variables, including the chemical composition and an array of rolling parameters [3]. Korczak *et al*. used microstructural parameters as inputs to calculate ferrite grain size and property distributions [4]. There are also many examples of modelling the mechanical properties of steels. Dumortier *et al*. have modelled the properties of microalloyed steels [5], Millytoski has published many papers about modelling various properties of steels including the hot torsion of austenite and a comparison of the models with physical models [6, 7].

The work presented here uses a neural network with a Bayesian framework. This approach was used to predict flow stress from the inputs of composition, temperature, strain rate and strain. The Bayesian approach to neural networks makes predictions with error bars, with the magnitude depending upon the position in the input space and perceived level of noise in the model. This should be extremely useful from an industrial point of view. Often, when data are needed to optimise processing schedules of a new composition of steel, the data are not readily available, and it is unclear to what extent data from related steels may be exploited. As a result constitutive data often need to be generated by testing for each new composition of steel.

Hot Deformation of Austenite

Hot−working often refers to deformation carried out under conditions of temperatures and strain rate such that recovery and recrystallisation processes occur substantially during deformation, achieving large strains with essentially no strain hardening. Hot−working processes such as rolling are typically the first step in converting a cast ingot into a wrought product such as steel strip. Hot−working is usually carried out at a homologous temperature of 0.6 of the melting temperature and at strain rates between 0.5 and 500 s^{-1} [8]. Laboratory tests for studying metallurgical changes during hot−working are either usually hot−torsion or compression tests.

The general behaviour of a wide range of materials in response to increasing strain at high temperatures is rising stress to a maximum, followed by decrease to a steady−state flow stress, as shown in figure 1.

There are two mechanisms responsible for softening in hot−working, depending on the metal [9]. In aluminium and alpha iron dynamic recovery is the softening mechanism. This occurs by the formation of a well−developed sub−grain structure by cross slip and climb, as occurs in creep deformation, and the activation energy for hot−working is that for creep and self−diffusion. In metals with a lower stacking fault−energy, the softening in hot working is higher than for creep, and the softening occurs by the mechanism of dynamic recrystallisation. The difference in the stress/strain curves for the two types of materials can be seen in figure 1.

Group	***SFE***	***Dynamic***	***Static***
Al, alpha−Fe, ferritic alloys	Low Stacking fault energy	Recovery at all strains	Recovery followed by recrystallisation
Cu, Ni, gamma−Fe, austenitic alloys	High Stacking fault energy	Recovery at small strains, Recrystallisation at high strains	Very limited recovery followed by recrystallisation

Table 1 Restoration Processes associated with hot working.

Dynamic recovery is the basic mechanism that leads to the annihilation of dislocations. It results in a flow curve about one order of magnitude lower than in cold working [8]. The flow stress/plastic strain curve is essentially exponential, the stress rising to a steady state value when the work hardening and softening mechanisms are in equilibrium at higher strains. Low dislocation densities associated with the deformation are due to the ease of cross slip, climb, and dislocation unpinning.

In metals that exhibit dynamic recrystallisation, dislocation annihilation only occurs when the dislocation density reaches high enough levels for recrystallisation to occur. As a result the flow stress rises to a peak value before dropping down to a steady state value where recovery and work hardening and in equilibrium.

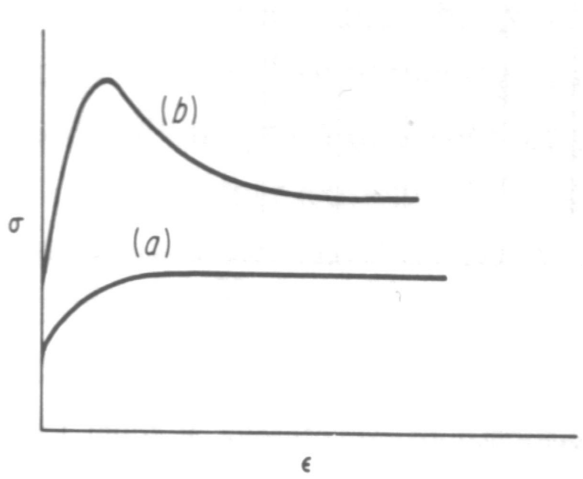


Figure 1 (a) Stress−strain curve for metal exhibiting dynamic recovery only; (b) metal which exhibits dynamic recrystallisation after an initial period of dynamic recovery.

Stress/Strain relationship

An elementary form of constitutive equation in common use is: [8] [10]

$$\sigma = K\varepsilon^{n} \tag{2.1}$$

where σ is the flow stress, ε is the plastic strain and *k* and *n* are fitting constants. The parameter *n* is often identified with strain hardening behaviour. It follows that:

$$n = \frac{\varepsilon}{\sigma}\frac{d\sigma}{d\varepsilon} = \frac{d(\log\sigma)}{d(\log\varepsilon)} \tag{2.2}$$

Other empirical relationships are sometimes used because equation 2.1 may not adequately fit the data, for example: [8]

$$\sigma = K(\varepsilon_0 + \varepsilon)^{n} \tag{3.1}$$

$$\sigma = K\varepsilon^{n} + \exp(K_1)\exp(n_1\varepsilon) \tag{3.2}$$

and also the Ludwik equation: [8]

$$\sigma = \sigma_0 + K\varepsilon^{n} \tag{3.3}$$

where σ is the flow stress, ε is the plastic strain and *k* and *n* are fitting constants as in equation 2.1. In a review paper on hot workability, Sellars and Teggart reported that the

following equations have in the past also been used to fit measured stress/strain curves: [19]

$$\sigma = A(B+\varepsilon)^m \quad \textbf{4.1}$$

$$\sigma = A + B\ln\varepsilon \quad \textbf{4.2}$$

$$\sigma = A - (A-B)\exp(-C\varepsilon) \quad \textbf{4.3}$$

$$\sigma = \sigma_0 + B\exp[1-\exp(-C\varepsilon)]^m \quad \textbf{4.4}$$

$$\sigma = A[1-\exp(-C\varepsilon^{mn})]^{\frac{1}{n}} \quad \textbf{4.5}$$

where σ is the flow stress, ε is the plastic strain and *A, B, m* and *n* are fitting constants, and σ_0 is the stress at the onset of flow.

One particular problem is that the derived parameters can rarely be used for alternative experimental data, and when they are used, the danger in doing so is not explicitly expressed. The constants are generally fitted using linear regression; an advantage in using the neural network technique over linear regression is that it does not rely on choosing a relationship before the analysis and is able to find non−linear relationships in the data.

Work Hardening

Plastic flow occurs by the mechanism of dislocations, the work hardening therefore is caused by the fact that dislocations are harder to generate or dislocations are harder to move. The ease with which dislocations are normally generated across a range of conditions suggests that increased resistance to dislocation movement causes the work hardening effect. This assumption is the basis of theories of work hardening.

Taylor [11] suggested that work hardening was due to the interaction of dislocations some of which become stuck and act as sources of internal stress which oppose the motion of other dislocations. Taylor suggests that for work hardening due to dislocation interactions, the flow stress to move the dislocations in the stress field is give by:

$$\sigma = \frac{aGb}{l} \quad \textbf{5.1}$$

This gives the force needed to force two dislocations onto parallel slip planes of spacing *l* against their elastic interaction. a is a constant in the Taylor model with a value of approx 0.1, given by:

$$a = \frac{1}{(8\pi(1-\upsilon))} \quad \textbf{5.2}$$

where ? is poisons ratio, b is the Burgers vector, G is the shear modulus and *l* is the mean distance between dislocations, which can be approximated to the square root of the dislocation density.

The macroscopic plastic shear strain can is given by: [12]

$$\varepsilon = b\rho_m \overline{x}$$

5.3

from equations 5.1 and 5.3:

$$\sigma = aG\left(\frac{b}{x}\right)^{\frac{1}{2}} \varepsilon^n$$

5.4

where n is 0.5 in the Taylor theory. Such a parabolic relation is seen in many polycrystalline materials. This does not entirely explain the behaviour observed in the stress−strain curve, with barriers to dislocation movement actually being created by the deformation.

Temperature and Strain rate dependence

The energy for dislocations to overcome the barriers they encounter during slip determines the dependence of the flow stress on the temperature and the applied strain rate. If the barriers are sufficiently small for the thermal energy to have a significant effect then thermal vibrations of the atoms may assist the dislocations to overcome obstacles at lower values of applied stress.

Consider a dislocation gliding in the *x* direction under the force of an applied resolved shear stress *?**, which produces a force of *?****b*** per unit length. Obstacles are separated by a distance *l* and resist the dislocation movement with a force of *K*. To move from x_1 to x_2 the isothermal energy required is given by the area under the *K* versus *x* curve between x_1 and x_2. The Helmholtz free energy change to overcome a barrier between x_1 to x_2 is: [12]

$$\Delta F^* = \int_{x_1}^{x_2} K.dx$$

6.1

Part of the energy required to overcome and obstacle to dislocation motion can be done by the applied load. The remainder or the energy required can be supplied thermally and is known as the *free energy of activation* [12].

$$\Delta G^* = \Delta F^* - \tau^* V^*$$

6.2

The probability of energy ?G occurring by thermal fluctuations at temperature *T* (such that *?G > kT*) is given by the Boltzman factor exp(−*?G* /kT*), so that the dislocation is effectively vibrating at a frequency *?*, overcoming *?* exp(−*?G* /kT*) barriers per second. This dislocation velocity is therefore.

$$\bar{v} = d\nu \exp\left(\frac{-\Delta G^*}{kT}\right)$$

6.3

The critically resolved shear stress resolved on the slip plane is give by:

$$\tau = \frac{F}{A}\cos\phi\cos\lambda$$

6.4

Thus from equations 6.3 and 6.4 the macroscopic plastic strain rate is: [12]

$$\dot{\varepsilon} = \rho_m A \exp\left(\frac{-\Delta G^*}{kT}\right)$$

6.5

where $?_m$ is the mobile dislocation density and $A=\boldsymbol{b}d?$. The stress dependence of the strain rate arises from the stress dependence of $?G^*$.

However the flow stress consists of two contributions. A thermal component, from short range interactions that can be overcome by thermal activations and a athermal component, from longer range interactions, which cannot be overcome by thermal activation. [12]

$$\tau = \tau^* + \tau_G$$

6.6

The temperature dependence of the critically resolved shear stress, and hence the flow stress, therefore depends upon the nature of the obstacles to dislocation movement. An increase in temperature or a decrease in applied strain rate provides an increase in the probability of thermal activation and therefore results in a reduction in flow stress.

Equation 6.5 demonstrates the equivalence between an increase in strain rate and an increase in temperature. In many models of hot working this equivalence is expressed as Z the Zener−Hollomon parameter. Zener and Hollomon [13], tested the equivalence of changes in strain rate and temperature. Proposing that an isothermal stress strain relationship depends upon strain rate and temperature only through a single parameter.

$$z = \dot{\varepsilon}\exp\left(\frac{Q}{RT}\right)$$

6.7

Effect of grain size

From experimental measurement of the yield stress of polycrystalline aggregates in which grain size is the only variable, it has been found that the Hall−Petch relationship is satisfied:

$$\sigma_y = \sigma_0 + k_y d^{-n}$$

7.1

where n is approximately 0.5, k_y is a material constant and $?_o$ is a constant stress.

Effect of alloying elements

The addition of alloying elements to the steel will alter the hot strength of austenite in three main ways:

1) The element can dissolve interstitially or substitutionally to form a single phase solid solution so that the austenite is strengthened.

2) The elements can combine with carbon and nitrogen to form precipitates so that the austenite is strengthened by precipitation strengthening.

3) The elements may have limited solubility in austenite and form a closed gamma−loop, At hot working temperatures such steels may have a duplex gamma−alpha structure. The strength depending on the proportions of gamma and alpha phases present, strength therefore depending upon composition and temperature.

In general the strength of single phase austenite increases with increasing alloy content [14].

From dislocation theory the effect of solutes atoms has been shown to influence the strength by a factor proportional to the square root of the concentration. [12, 15]

Interstitial solutes

In the absence of strong carbide forming elements carbon dissolves interstitially in austenite, with the limit to solubility being up to 10 at% carbon [16]. Tegart suggested it should have little effect on the hot strength by solute strengthening because of its high diffusion rate [14]. Tegart observed that there was little increase in hot strength with carbon additions at temperatures above 900°C and that there was a tendency to decrease at higher carbon concentrations [14].

In highly alloyed austenite, carbon is reported to lead to decreased strength at higher concentrations while nitrogen increased strength. This difference has been related to the difference in stacking fault energy, which is relatively low in these alloys and increased by carbon and decreased by nitrogen.

It has also been reported that carbon increases hot strength in commercial steels. While Stewart reported that the flow stress in the austenite region is nearly independent of the carbon level [17].

Substitutional solutes

Systematic studies by Zidek [14] on binary alloys shows that some elements are more effective strengtheners than others. Nickel appeared to have no effect at either 9 wt% or 40 wt% whereas chromium, manganese, and molybdenum lead to substantial increases.

Elements such as molybdenum and manganese were reported by Tegart to have strengthening effects both low alloy steels and 18/8 stainless steel and suggested that the alloying elements strengthen austenite independently having an additive effect on the strength.

It should be noted that the relative strengthening effect may alter with the temperature of testing, and with the concentrations of interstitial solutes present in the steels. Table 2 shows the change in the change in hot strength due to several substitutional alloying elements at different hot working temperatures.

Element	**Atomic Diam. Å**	**% Diff in size relative to Fe**	**% change in hot strength for 1%**		
			1000°C	**1100°C**	**1200°C**
Si	2.36	−8.5	8.3	2.4	−3.6
Ni	2.48	−3.9	−0.1	−0.2	−0.3
Cr	2.51	−2.7	2.1	1.6	1.1
Mn	2.70	4.7	5.3	4	2.8
Mo	2.74	6.2	13	10.5	7.9
Al	2.95	15.9	16.3	13.6	9.5

Table 2, Strengthening effects of alloy additions **[14]**

The shear strength depends upon the size difference between the solute and solvent atoms, in general the change in strength is given by:

$$\frac{d\tau}{dc} \propto \left(\frac{1}{a} \frac{dc}{da} \right) \qquad \textbf{7.2}$$

where a is the lattice parameter of the solid solution, c is the solute concentration.

Precipitates

Precipitates give significant increases in hot strength of austenite when present as finely dispersed phases. These are usually carbides, nitrides or carbonitrides and are often present in steels that have also been strengthened by solid solution strengthening.

The kinetics of solution and precipitation of second phases are strongly temperature and time dependent, and also influenced by deformation, so varying results can be obtained with steel depending upon its temperature and deformation history.

Duplex structures

The strength of duplex structures does not vary linearly with the proportion of phases but also the strength of continuous phase. The alloy content is of secondary importance and the strength can even decrease if an alloy addition such as Cr increases the proportion of

alpha−phase present due to the low strength of ferritic steels compared to austenitic steels under hot working conditions.

Effect on restoration Processes

The constitution and microstructure has an effect on the movement of dislocations and also the thermal activation of dislocation movement.

Colàs reported the activation energies reported for steels varied between 230 and 390 k $Jmol^{-1}$. Using linear regression on low alloy steels and assuming the activation energy is independent of microstructure, the activation energy, Q, in $Jmol^{-1}$ was found to be given by:

$$Q = 282700 + 92400(\%C) + 6570(\%Mn) + 1000(\%Si) \quad \textbf{7.3}$$

were the compositions are in wt%. Valid for chemical compositions ranging from 0.03−0.3 wt% carbon, 0.3−1.7 wt% manganese and 0−0.6wt% silicon. Medina and Hernandez also carried out multiple regression analysis, to obtain:

$$\begin{aligned} Q = \quad & 267000 - 2535.52(\%C) + 1010(\%Mn) + 33620.76(\%Si) + 35651(\%Mo) \\ & + 93680.52(\%Ti)^{0.5919} + 31673.46(\%V) + 70729.85(\%Nb)^{0.5649} \end{aligned} \quad \textbf{7.4}$$

Medina and Hernandez reported that the peak stress in niobium steels was not much higher than in carbon−manganese steels, and the value of peak stress in vanadium steels was also similar at strain rates of 0.5 to 5 s^{-1}. They made the conclusion that neither niobium nor vanadium precipitated during deformation and the strain is much less than the incubation time for dynamic precipitation. They also observed that all the alloying elements increased the peak stress, apart from carbon, which was found to have a slight softening effect [18].

Constitutive Equations

Early experimental work of Hollomon [10] also showed the applicability of the mechanical equation of state, where the stress required for flow depends upon the instantaneous values of strain, strain rate and temperature. The stress under certain conditions does not depend upon the past history, but is analogous to the temperature of a gas, which depends only upon the pressure and volume.

$$\sigma = f(\varepsilon, \dot{\varepsilon}, T) \tag{8.1}$$

The temperature and stress dependence of the steady−state hot working rate is often expressed by the empirical relationship: [8, 18]

$$\dot{\varepsilon}_s = A(\sinh \alpha\sigma)^{n'} \exp\left(\frac{-Q}{RT}\right) \tag{8.2}$$

Where $\dot{\varepsilon}$ is the steady−state strain rate, σ is stress, A, α and n' are experimentally determined constants. At low stresses equation 8.2 reduces to a power relation that describes creep behaviour.

$$\dot{\varepsilon} = A_1 \sigma^{n'} \exp\left(\frac{-Q}{RT}\right) \tag{8.3}$$

While at high stresses equation 8.2 reduces to an exponential form

$$\dot{\varepsilon} = A_2 \exp_{10}(\beta\sigma) \exp\left(\frac{-Q}{RT}\right) \tag{8.4}$$

Equations for strain and strain rate dependence can be combined to give equations of the form: [19]

$$\sigma = (BZ)^{\frac{1}{n}} \varepsilon^m \tag{8.5}$$

$$\sigma = \frac{\varepsilon^m}{\beta} \ln BZ \tag{8.6}$$

$$\sigma = \frac{\varepsilon^m}{\alpha} \sinh^{-1}(BZ)^{\frac{1}{n}} \tag{8.7}$$

While the limits of the algebraic equations is recognized, the relationships are widely used in reporting experimental data in the form of the constants for a particular material rather then publishing complete stress/strain curves, from which flow stress or work done for any strain can be computed.

Plastic deformation is an irreversible process; the general form of the constitutive equation is,

$$d\sigma = \left(\frac{\partial \sigma}{\partial \varepsilon}\right)_{\dot{\varepsilon},T} d\varepsilon + \left(\frac{\partial \sigma}{\partial \dot{\varepsilon}}\right)_{\varepsilon,T} d\dot{\varepsilon} + \left(\frac{\partial \sigma}{\partial T}\right)_{\varepsilon,\dot{\varepsilon}} dT \qquad \textbf{8.8}$$

A form of which has been shown to be valid at large strains by Ghosh [20] is,

$$\frac{\dot{\sigma}}{\sigma} = \dot{\varepsilon}\gamma + \left(\frac{\ddot{\varepsilon}}{\dot{\varepsilon}}\right)m + \dot{T}\frac{(\partial \sigma / \partial)}{\sigma}$$

where
$\dot{\sigma} = \partial\sigma / \partial t$
$\dot{\varepsilon} = \partial\varepsilon / \partial t$
$\ddot{\varepsilon} = \partial\dot{\varepsilon} / \partial t$
$\gamma = (\partial\sigma / \partial\varepsilon)/\sigma$
$m = (\partial\dot{\varepsilon} / \partial\sigma)(\partial\sigma / \partial\dot{\varepsilon})$
$\dot{T} = \partial T / \partial t$

8.9

Similar forms are used in computer representations such as the MATMOD code for high temperature deformation [21].

Hartley and Srinivasan developed another form of equation to describe the stress−strain relation [22]. For an isothermal constant true strain rate test, equation 8.9 can be simplified to:

$$\gamma \propto (\varepsilon + \varepsilon_0)^{n-1} \qquad \textbf{8.10}$$

or

$$\gamma \propto \left(\frac{\sigma_s - \sigma}{\sigma}\right)(\varepsilon + \varepsilon_i)^{r-1} \qquad \textbf{8.11}$$

on integration these give the saturation type equations:

$$\sigma = \sigma_0 \exp\left(\left(\frac{\varepsilon + \varepsilon_0}{\varepsilon^*}\right)^n\right) \qquad \textbf{8.12}$$

and:

$$\sigma = \sigma_s - (\sigma_s - \sigma_i)\exp\left(-\left(\frac{\varepsilon + \varepsilon_i}{\varepsilon^*}\right)^r\right) \qquad \textbf{8.13}$$

Neural Networks

Neural networks are statistical models of real world systems, built by tuning a set of parameters known as weights. The weights make up a model, which represents a mapping from the input values to the output values. The weights are calculated by passing examples of input−output pairs through the model, and adjusting the weights to minimise the error or prediction, with appropriate measures to avoid overfitting.

There are two major tasks that neural networks can be applied to: classification and continuous numeric functions. Classification refers to variables which take in a 0 or 1 value. It is continuous numeric functions that are of the most relevance to modelling constitutive behaviour.

Neural networks represent a general method of regression that can overcome some of the difficulties associated with ordinary linear regression: such as the need to choose a relationship between the parameters before analysis. The artificial neural network arrives at a mathematical model without prior assumptions about the form of the relationships. Relationships in the neural network are not limited to the sum of linear and pseudo−linear terms, and because of the non−linearity can be difficult in different regions of the input space.

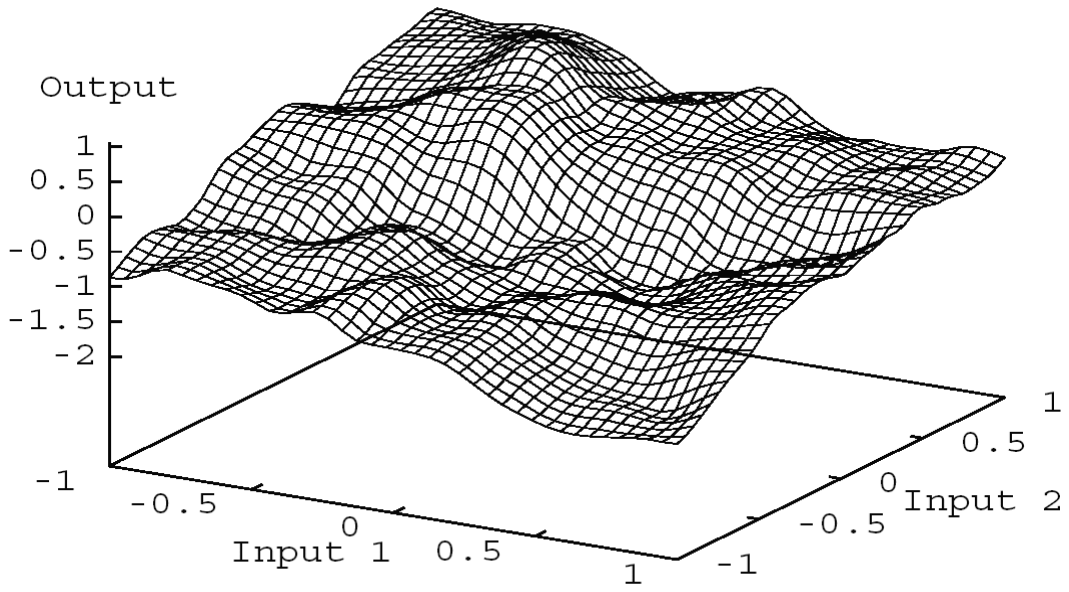


Figure 2, Typical function produced by neural network. **[23]**

Bayesian probability theory provides a unifying framework for data modelling which offers several benefits. Overfitting can be avoided by using methods to control model complexity, while probabilistic modelling handles uncertainty in a natural manner [23].

Using neural networks within a Bayesian framework allows uncertainties of fitting to be estimated in a manner which depends upon the region of the input space where the prediction is calculated. Instead of calculating a unique set of weights, a probability distribution of sets of weights is used to define the fitting uncertainty. This methodology

is extremely useful when applied to problems in materials science where properties need to be estimated as a function of a large number of inputs, which are not uniformly distributed in the input space [2].

Neural networks can create functions with much more flexibility than ordinary linear regression. Figure 2, shows a typical function produced using a neural network with two inputs and one output; in contrast the function produced by a linear regression would be a flat plane [23].

The final output, *y*, is defined as:

$$y = \sum w_i^{(2)} h_i + \theta^{(2)}$$

9.1a

where,

$$h_i = \tanh\left(\sum_j w_{ij}^{(2)} x_j + \theta_i^{(1)} \right)$$

9.1b

Where x_j, are the *j* variables on which the output *y* depends, w_i are the weights and $?_i$ are the biases.

The form of the relationship described by equation 9.1 is interesting. A hyperbolic tangent function (eqn. 9.1b is used to operate the weights inputs because such a function is non−linear and flexible in the sense that its shape is dependent on the weights. Combining several hyperbolic tangents together gives even greater flexibility so that the complexity of the model is related to the number of hyperbolic tangent operators used.

However the tanh function has a range from ±1 whereas the output can in principle range from ±?. The linear operator in equation 9.2 is enables the range to exceed ±1.

The input data are also scaled using equation 9.2, although this is not strictly necessary for fitting, it allows better comparison between the effects of different variables.

$$x_n = \left(\frac{x - x_{\min}}{x_{\max} - x_{\min}} \right) - 0.5$$

9.2

Using the normalised data the weights and bias were minimised in such a way as to minimise the objective function;

$$M(w) = \beta E_D + \sum_C \alpha_C E_{W(C)}$$

9.3

The minimisation was implemented using a variable metric optimiser. The gradient of M_w was computed using back−propagation algorithm. The objective function or energy function consists of an error function, E_D and regularisation E_W. The error function is the sum squared error as follows;

$$E_D(w) = \frac{1}{2}\sum_m \left(y(x^m; w) - t^m\right)^2$$

9.4

where $\{x^m, t^m\}$ is the data set, where *x* represents the inputs and *t* the targets and *m* is the number of the input− target pair. The error function is smallest when the model fits the data well. A number of regularization terms are added to the data error. These favour functions which are smooth functions of *x*. The simplest regularisation method uses a single regulariser.

$$E_w = \frac{1}{2}\sum w_i^2$$

9.5

In the work presented the method of regularisation was the automatic relevance determination model. Each weight is assigned to a class, *c*, depending upon which neurons it connects. For each input, all the weights connecting that input to the hidden units are in a single class, and all the weights from the hidden units to the outputs are in the final class. $E_{w(c)}$ is defined as the sum of the squares of the weights in class *c*.

$$E_{w(c)}(w) = \frac{1}{2}\sum_{i\in c} w_i^2$$

9.6

This additional term favours small values of the weights and decreases the tendency of the model to overfit noise in the data set. The control parameters alpha and beta together with the number of hidden units determine the complexity of the model. These hyper− parameters define the assumed Gaussian noise level.

$$\sigma_v^2 = 1/\beta$$

9.7

and the assumed weight variances

$$\sigma_{w(c)}^2 = 1/\alpha_{(c)}$$

9.8

$?_v$ is the noise level inferred by the model. The hyperparameter alpha has the effect of encouraging the weights to decay. A high value of $?_w$ associated with an input parameter implies that the input explains a relatively large amount of variation in the output. $?_w$ is then regarded as good expression of the significance but not the sensitivity of the output to the input.

Many equally valid models can be produced to model the same data The approach used here creates many models by starting training with different numbers of hidden units and starting values of $?_w$, the models can then be assessed by the test error (sum squared error) associated with each. It is then possible to reduce the test error further by using the average predictions from a number of models. The mean prediction given by:

$$\bar{y} = \frac{1}{N}\sum_{i-1}^{N} y_i$$

9.9

And the associated error in $\bar{y}$ given by:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 + \frac{1}{N}(y_i - \bar{y})^2$$

9.10

Without changing the complexity of the sub−models the committee model once selected was then retrained using the whole database.

The implicit assumptions of the Bayesian approach, proposed by MacKay [23], used here are that; the target is a noisy version of the true output, the noise can be estimated by a Guassian distribution with at variance that is the same for all examples, and the input variables are noise−free.

Neural networks to predict constitutive behaviour

Previous work has demonstrated the ability of a neural network in modelling constitutive behaviour. Narayan *et al.* [25] has demonstrated that hot− torsion stress/strain curves can be represented using a neural network technique, even taking into account the deformation history of the material.

Hwu, *et al.* [26] developed a neural network for prediction of flow stress using data from six steels, they reported that their neural network model could interpolate flow stresses very well but that the capability for extrapolation was not impressive. An attempt was made to build results from linear regression for carbon equivalence into the model and they found that this lowered the quality of the training.

Kong and Hodgson and co−workers have integrated a constitutive and neural network models reporting significantly improved the prediction accuracy with changes in chemical composition. [27] The integrated model they developed predicted the parameters of the Estrin−Mecking model from the input parameters of carbon content, temperature, strain rate, Zener−Hollomon parameter and activation energy.

Liu *et al.* [28] compared using neural networks with the use of the Zener−Hollomon parameter and hyperbolic sine stress function to model the behaviour of high−speed steel.

Experimental detail

Data was provided by Corus' Swindon Technical centre included constitutive data describing the behaviours of low, medium, and high carbon steels, carbon−manganese steels, High Strength low alloy steels, and austenitic steels.

The data took the form of stress strain data, typically the tests were up to 0.7, and temperatures of hot working 700−1200°C, strain rates varying from 0.01 to 100 s^{-1}. The compositions of the 24 steels used in training the model are listed in table 3 below, the distribution of the data is shown in figures 3.1−3.11. The amount of information about the compositions of each alloy varied, since the data was collected from various sources, rather than being generated by a systematic test program designed specifically for the purposes of neural network analysis. The data is from commercial grades of steel, rather than alloys designed to extract information about the physical behaviour of steels, with variation of composition.

Data for a further 6 steels were collected from published results. The compositions of these steels is shown in table 4 [17]. These results were used to demonstrate the predictive abilities of the models produced.

Alloy	Composition (wt%) balance Fe							
	C	Si	Mn	Cr	Ni	Mo	Cu	V
78C_railsteel	0.780	0.230	0.850	0.013	0.011	0.010	0.023	−
50Concast	0.180	0.300	1.320	0.015	0.019	0.001	0.011	0.080
55Concast	0.150	0.200	1.300	0.019	0.020	0.001	0.018	0.150
43Concast	0.200	0.240	1.210	0.022	0.023	0.004	0.023	0.001
100C	1.000	0.190	0.170	0.100	0.090	−	−	−
55C	0.560	0.360	0.280	0.120	0.090	−	−	−
15C_A	0.147	0.270	0.480	0.070	0.099	−	0.275	−
En16	0.350	0.270	1.490	0.030	0.110	0.280	−	−
En31	1.060	0.220	0.460	1.410	0.170	−	−	−
En52	0.470	3.740	0.580	8.200	0.200	−	−	−
VniCu	0.090	0.307	1.430	0.020	0.223	0.003	0.324	0.132
En45	0.610	0.940	1.580	0.120	0.270	0.060	−	−
En40	0.260	0.570	0.350	3.030	0.290	0.490	−	−
En25	0.350	0.270	0.660	0.590	2.540	0.590	−	−
304	0.070	0.430	0.480	18.600	7.700	−	−	−
316	0.070	0.670	1.340	17.290	12.040	2.260	−	−
20C	0.190	0.040	0.860	−	−	−	−	−
16C_V	0.160	0.260	1.220	−	−	−	−	0.064
95C	0.960	0.087	0.910	−	−	−	−	−
25C	0.250	0.080	0.450	−	−	−	−	−
12C	0.120	0.200	0.050	0.080	−	−	0.100	−
Leaded steel	0.078	0.003	1.220	−	−	−	−	−
08C	0.087	0.003	0.340	−	−	−	−	−
45C	0.430	0.260	0.740	−	−	−	−	−

Table 3, Chemical compositions of the alloys used to train the network

Alloy	Composition (wt%) balance Fe		
	C	Si	Mn
Steel1	0.010	0.150	0.190
Steel2	0.030	0.320	0.620
Steel3	0.190	0.310	0.640
Steel4	0.380	0.340	0.640
Steel5	0.590	0.640	0.350
Steel6	0.900	0.730	0.460

Table 4, Chemical compositions of the alloys used to show the predictive power of the models

It can be seen from looking at figures 3.1 to 3.11 that the data is not evenly distributed in the input space, but takes a series of discrete data points. The effect of the distribution of data points can be seen in the prediction of the error bars of the models presented later. Where no value was reported for an alloying element in a grade of steel, and were the alloying element is usually added on purpose to have a specific desired effect, rather than being a tramp element, the percentage was assumed to be 0. The distribution of many of the alloying elements was mainly concentrated at low percentages for example the chromium, and nickel which had a large range because of the inclusion of stainless steels in the database.

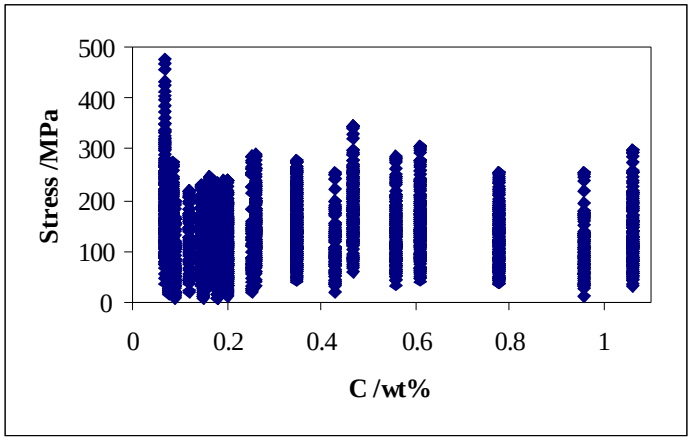


Figure 3.1, showing the distribution of carbon data

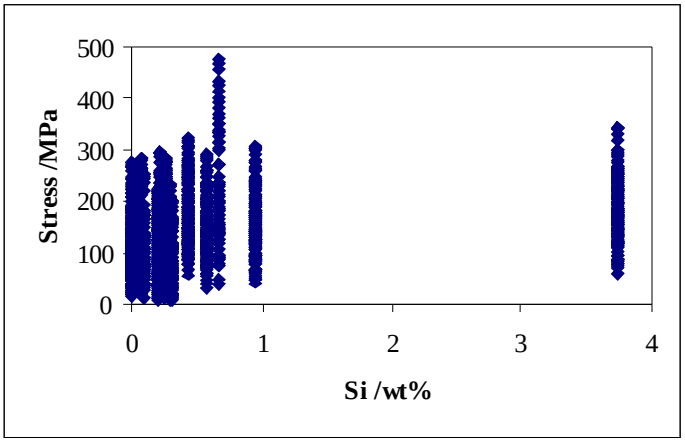


Figure 3.2, showing the distribution of silicon data

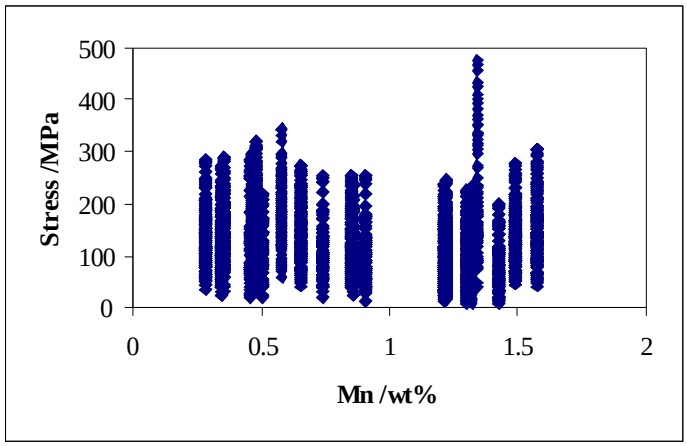


Figure 3.3, showing the distribution of manganese data

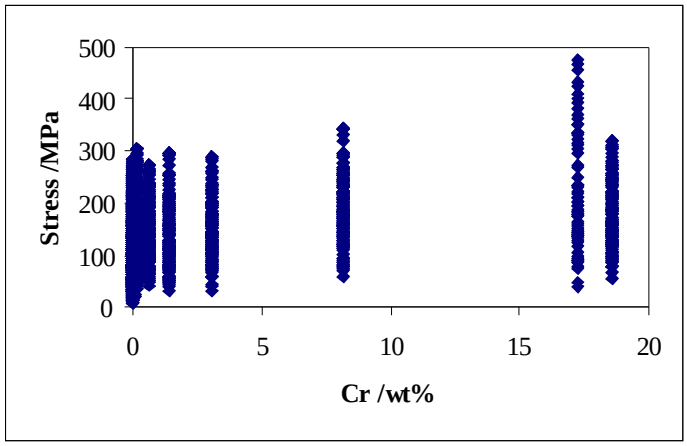


Figure 3.4, showing the distribution of chromium data

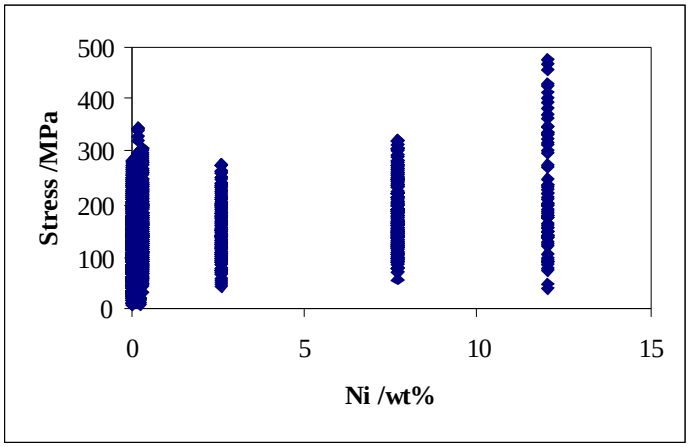


Figure 3.5, showing the distribution of manganese data

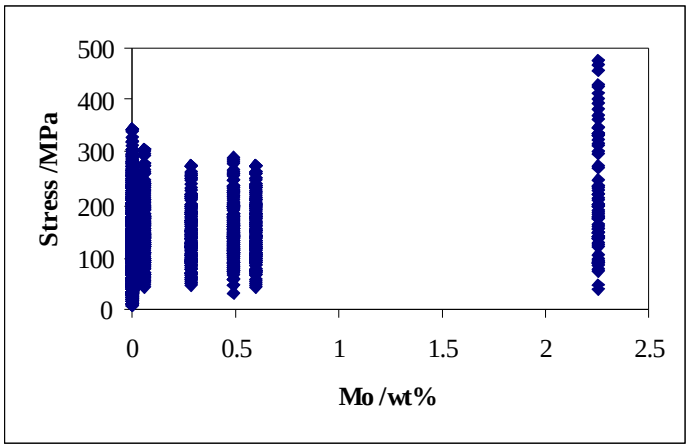


Figure 3.6, showing the distribution of molybdenum data

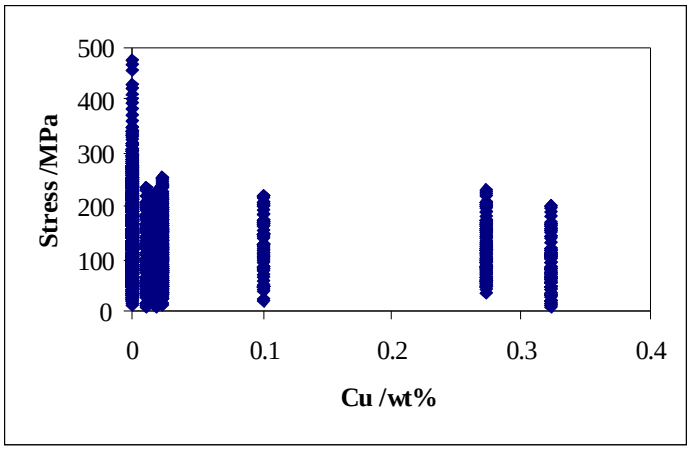


Figure 3.7, showing the distribution of copper data

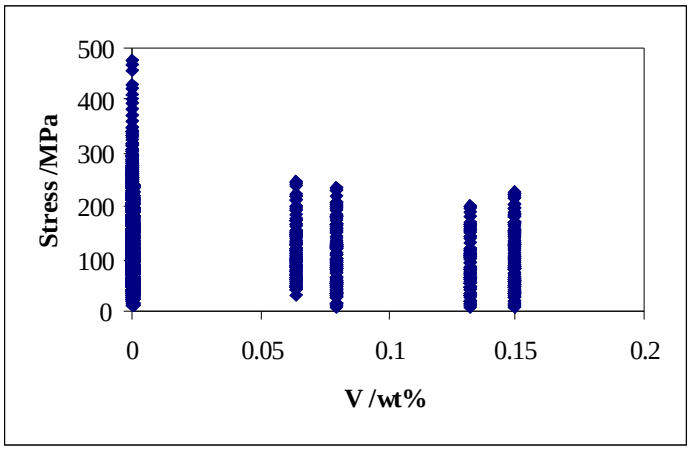


Figure 3.8, showing the distribution of vanadium data

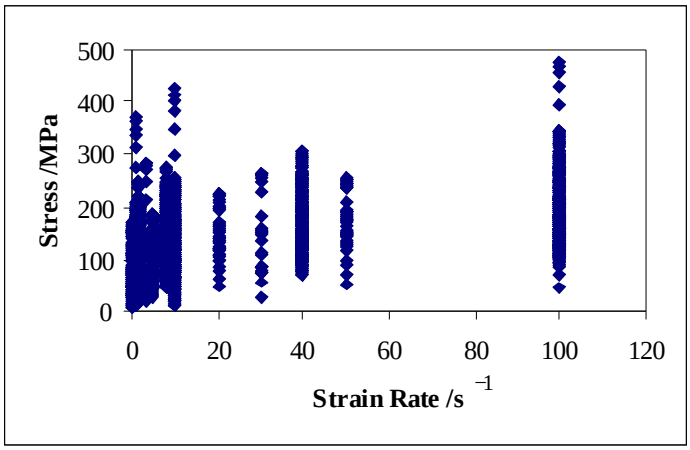


Figure 3.9, showing the distribution of strain rate data

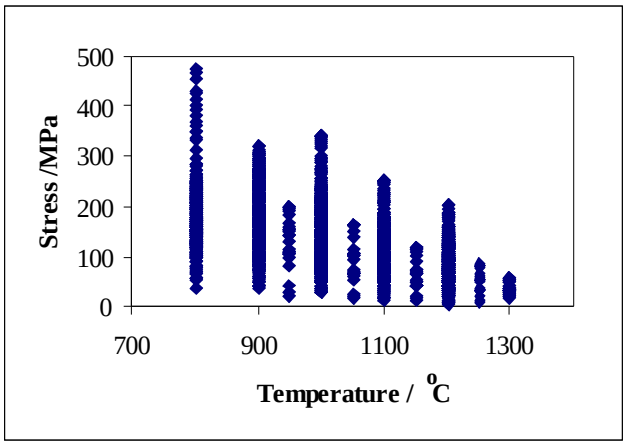


Figure 3.10, showing the distribution of temperature data

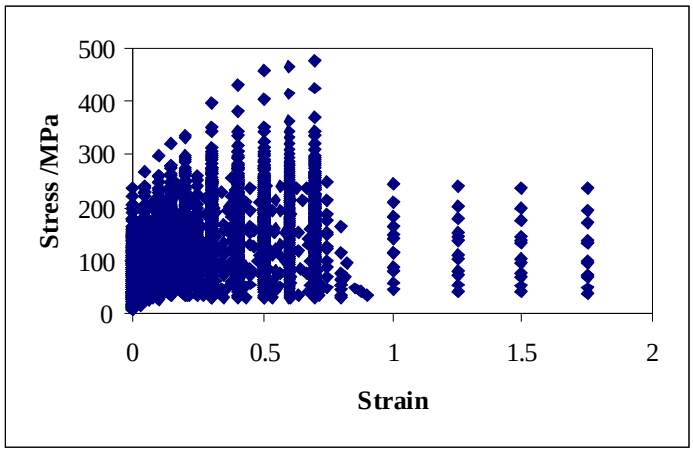


Figure 3.11, showing the distribution of strain data

The first model produced used the data as it was reported without trying to give the model any physical basis. One aim of this model is to look for trends in the data, another was to decide which data points should be included in the physical model. Phosphorus and sulphur data were available; these elements were left out to simplify the model since they are mainly tramp elements, with deleterious effects on properties of steel.

Table 5, The variables used in the neural network analysis, model 1.

Once the data has been collected and stored it is then normalised. Sub−models are selected and trained; these differ in the initial number of hidden units and starting weight values. The data was randomised and split into a training set and a testing set, figure 5 shows the performance of the committee and sub−models used in model 1.

Figure 4 shows the bar chart of $?_w$ the inputs in predicting the stress, model 1 using the two sub−models, which combined to give the lowest test error. The perceived level of noise in the model was $?_{nu}$ = 0.0181. A large value of $?_w$ implies that the input concerned explains a relatively large amount of the variation in the stress in the database.

Figure 6.1 − 6.11 show the sensitivity of individual variables for a particular position in the input space. It can be seen that increasing temperature causes a reduction in flow stress as expected from the dislocation theory and the Zener−Hollomon parameter discussed in the introduction, which suggests that the effect of temperature should be linear. The effect of carbon at the temperature and strain rate selected was to increase the flow stress.

Silicon and manganese reduced the flow stress this disagrees with the observations reported in the literature, particularly Medina and Hernendez who reported that carbon caused softening while the other elements increased strength.

Chromium and nickel had negligible effect on the flow stress. Molybdenum, vanadium and copper all increased the flow stress of the alloy Grade25C, tested at a temperature 800?C, strain rate of 30 s^{-1} and strain of 0.5.

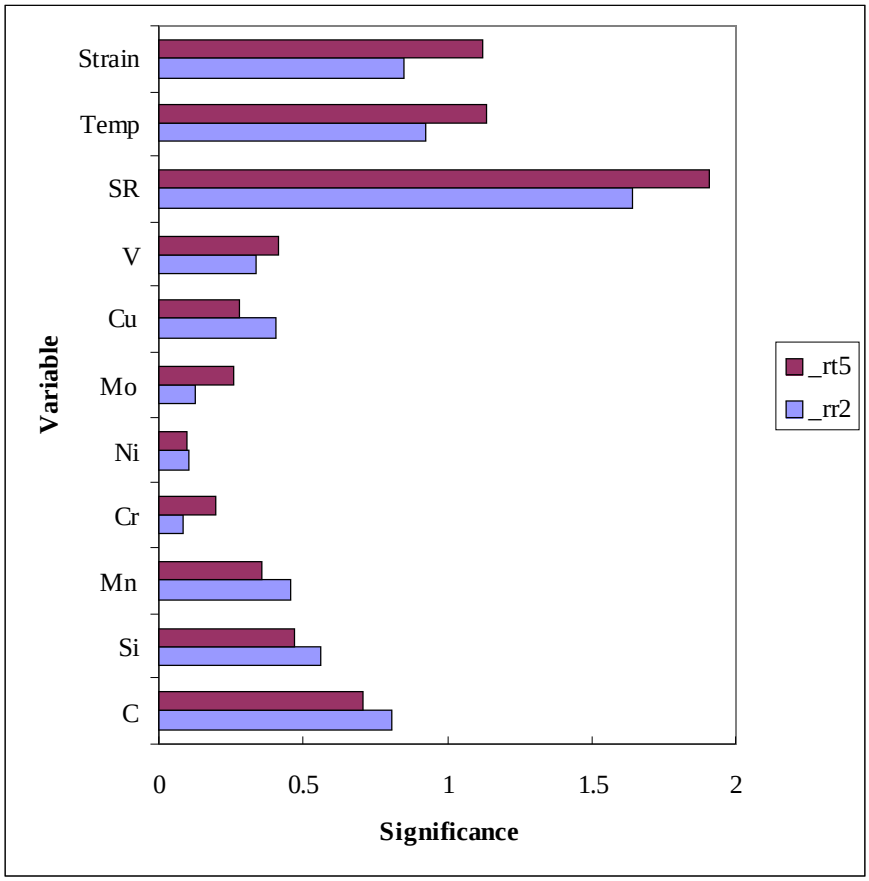


Figure 4, Bar−chart of the model perceived significance for each input for each sub−model in the committee, ?$_w$ values for model 1.

Figure 6.10, shows the effect of changing strain rate was to increase the flow stress, plotting ln(strain rate) against ln(stress) resulted in a straight line. Figure 6.11 compares the predicted stress−strain curve against the experimental values and shows good agreement, plotting ln(strain) against ln(stress) also resulted in a straight line.

Figure 5, training of model 1.

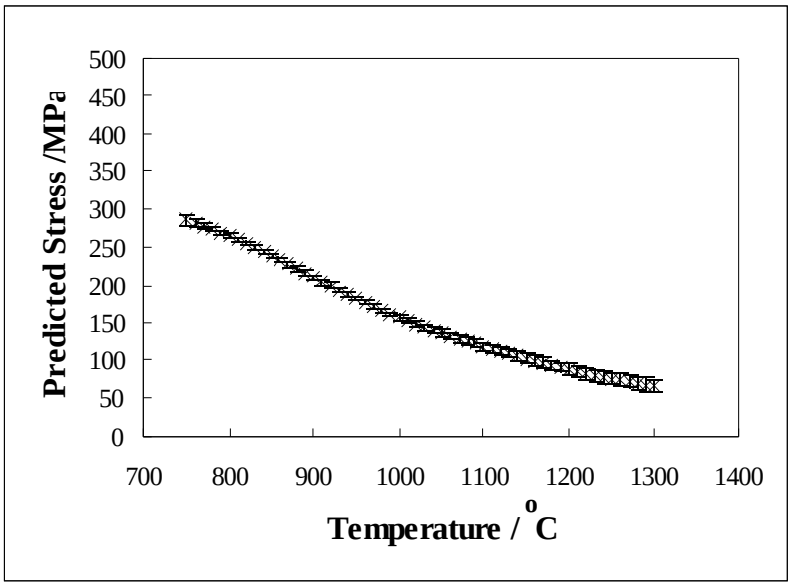


Figure 6.1, effect of temperature change on Grade25C, strain rate of 30 s^{-1} and strain of 0.5

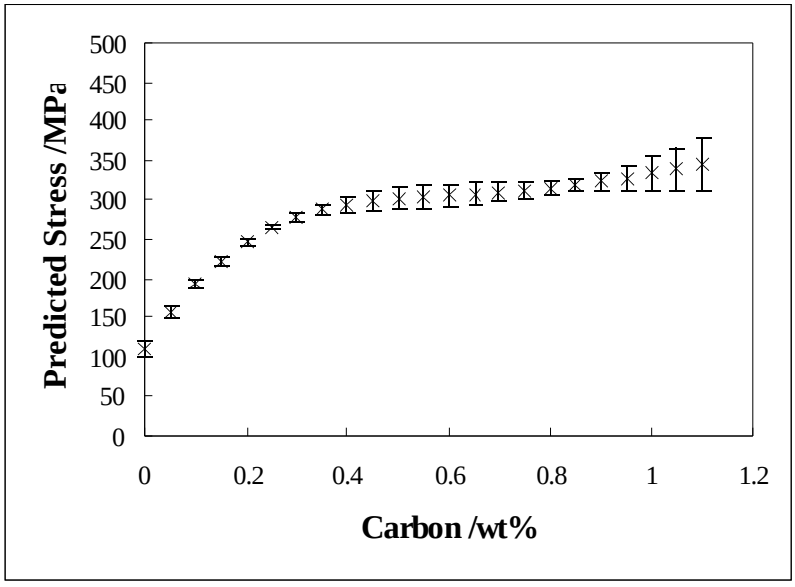


Figure 6.2, effect of carbon, Grade25C, temperature 800?C, strain rate of 30 s^{-1} and strain of 0.5

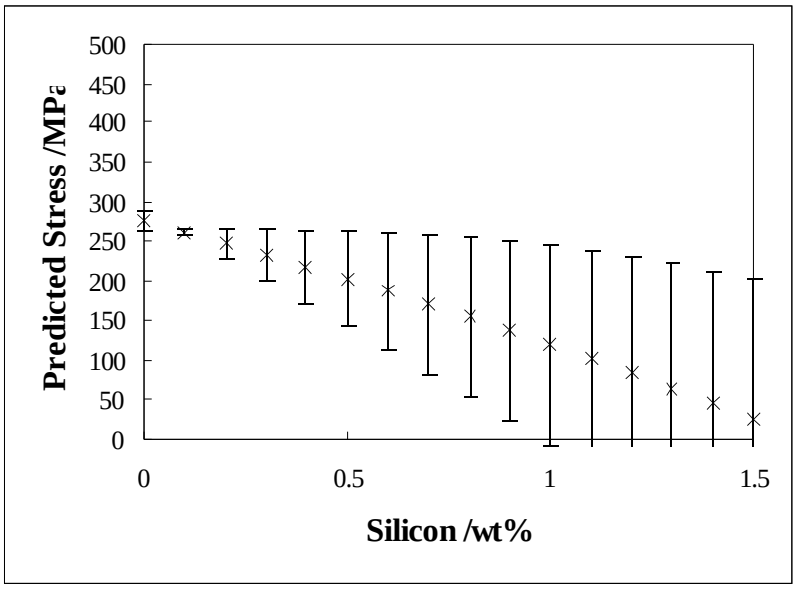


Figure 6.3, effect of silicon, Grade25C, temperature 800?C, strain rate of 30 s^{-1} and strain of 0.5

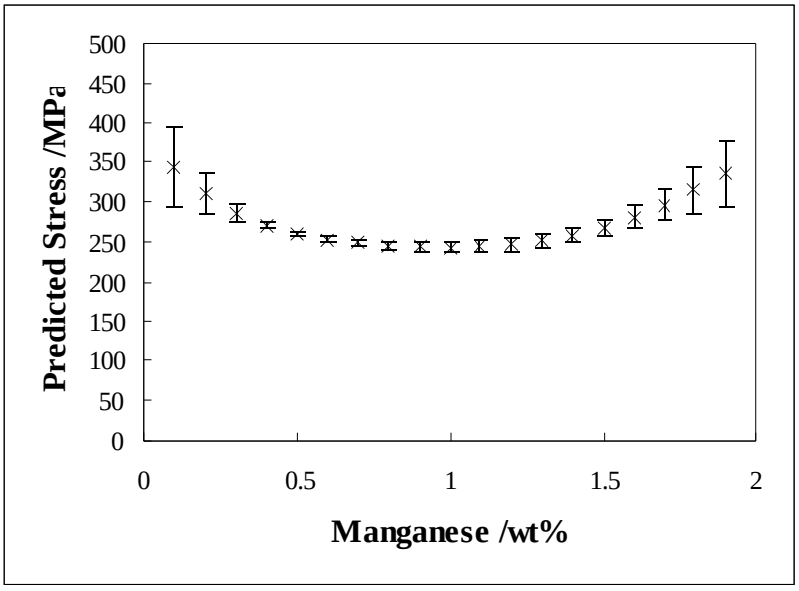


Figure 6.4, effect of manganese, Grade25C, temperature 800?C, strain rate of 30 s^{-1} and strain of 0.5

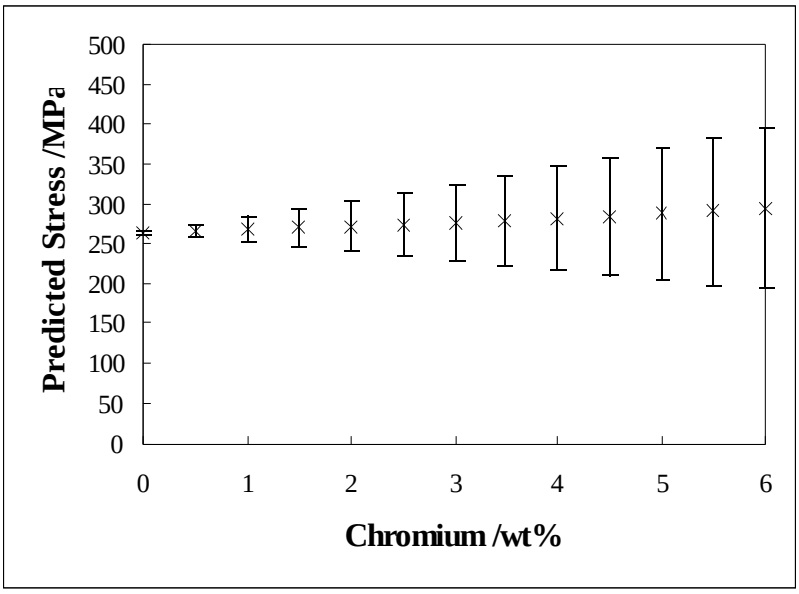


Figure 6.5, effect of chromium, Grade25C, temperature 800?C, strain rate of 30 s^{-1} and strain of 0.5

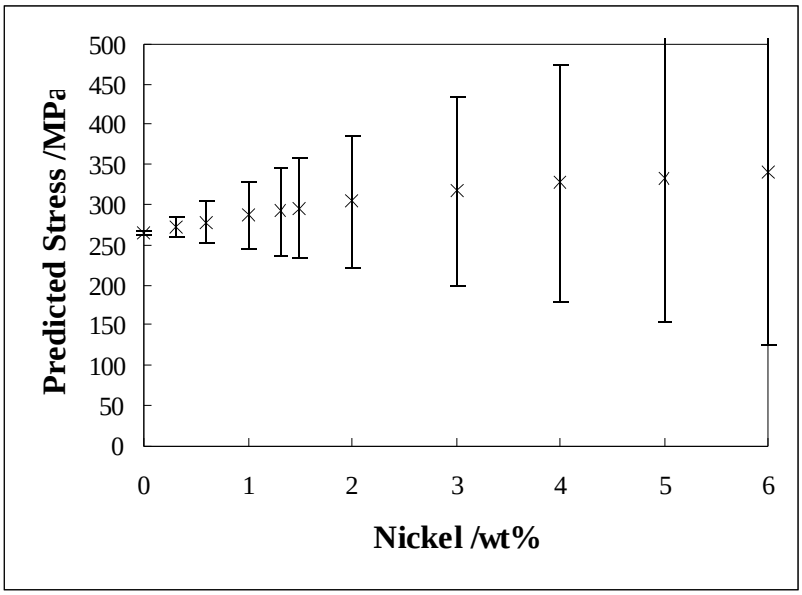


Figure 6.6, effect of nickel, Grade25C, temperature 800?C, strain rate of 30 s^{-1} and strain of 0.5

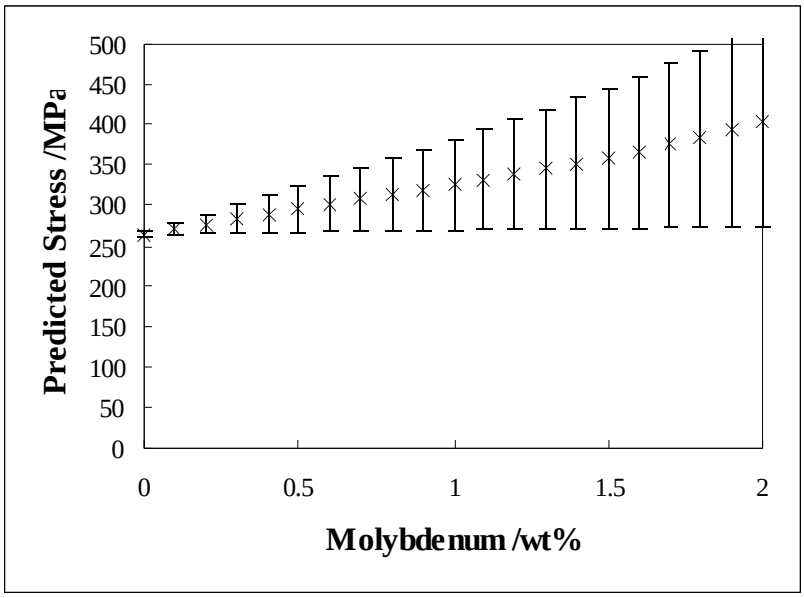


Figure 6.7, effect of nickel, Grade25C, temperature 800?C, strain rate of 30 s^{-1} and strain of 0.5

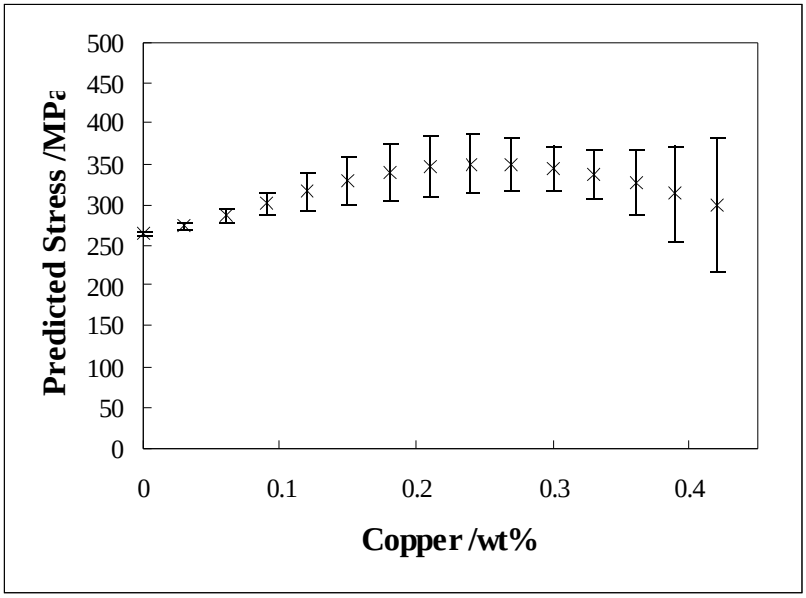


Figure 6.8, effect of copper, Grade25C, temperature 800?C, strain rate of 30 s^{-1} and strain of 0.5

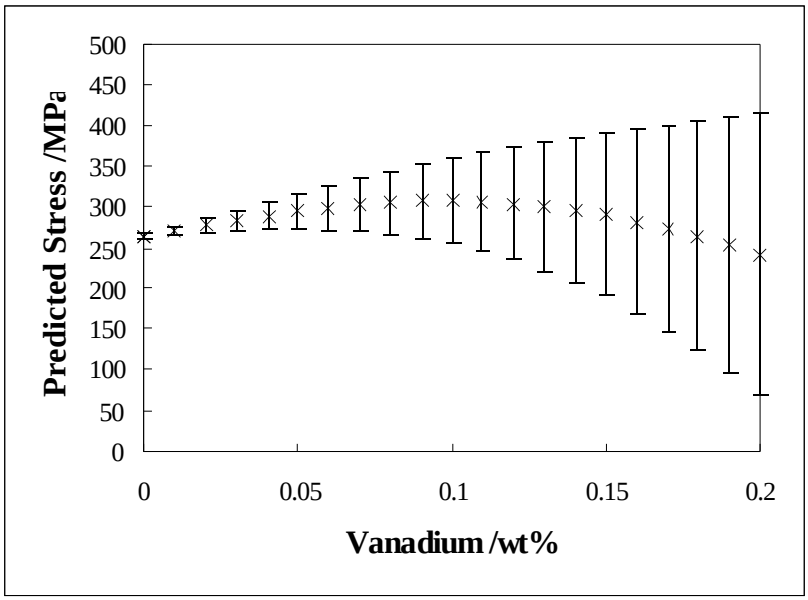


Figure 6.9, effect of vanadium, Grade25C, temperature 800?C, strain rate of 30 s^{-1} and strain of 0.5

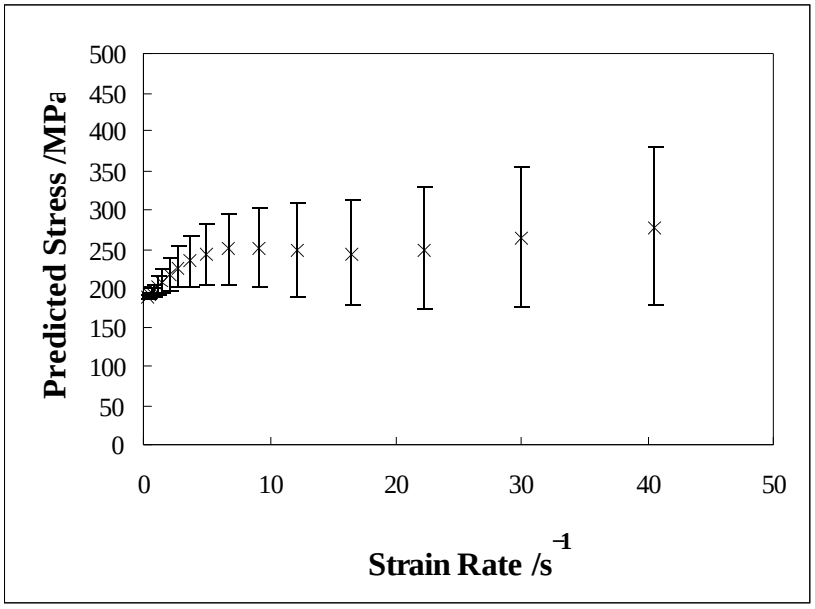


Figure 6.10a, effect of strain rate, Grade25C, temperature 800?C, and strain of 0.5

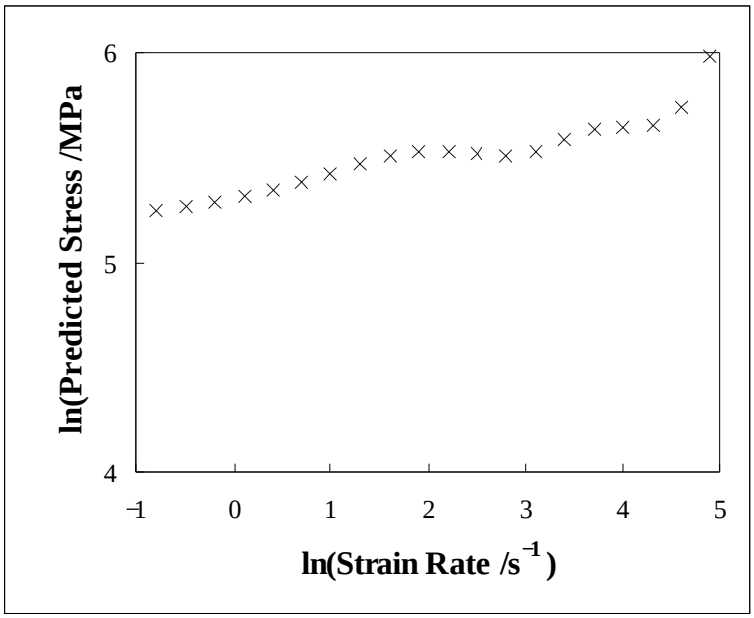


Figure 6.10b, effect of strain rate, Grade25C, temperature 800?C, and strain of 0.5

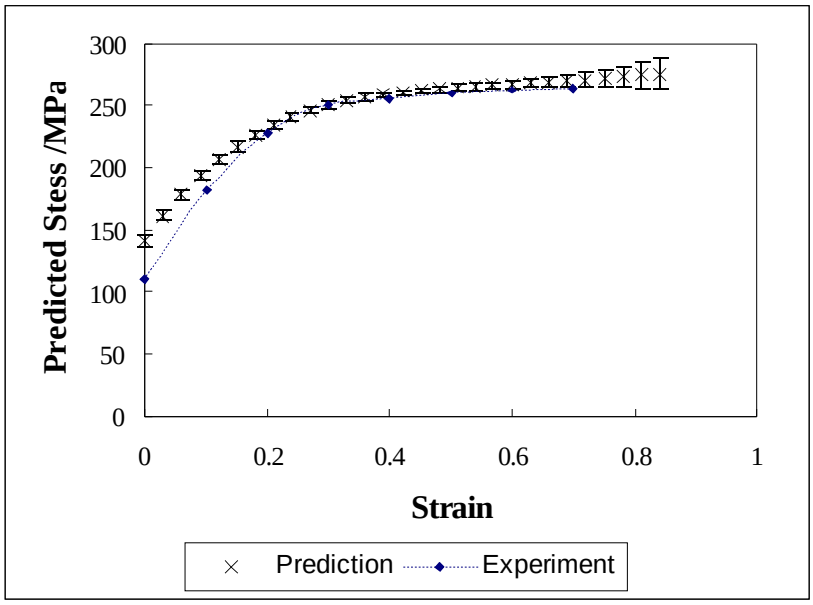


Figure 6.11a, effect of strain, Grade25C, temperature 800?C, and strain rate of 30s^{-1}

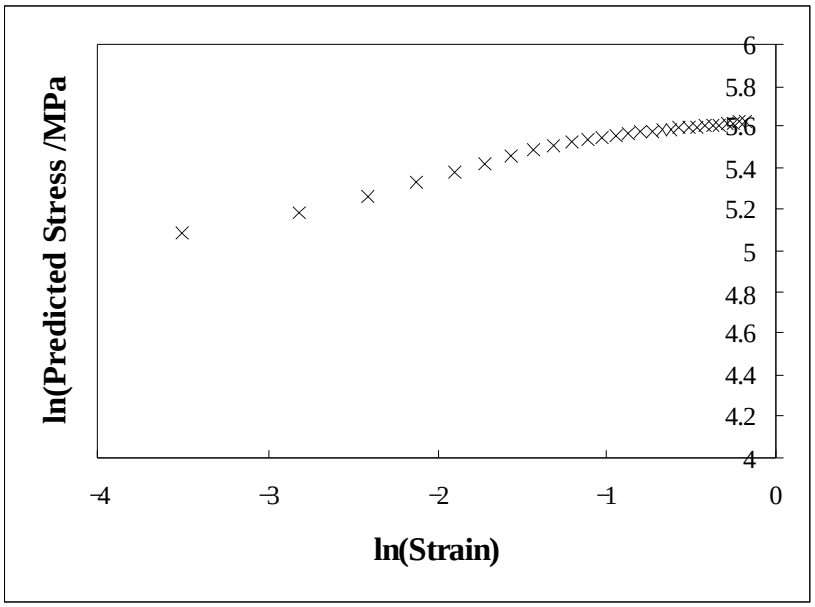


Figure 6.11b, effect of strain, Grade25C, temperature 800?C, and strain rate of 30 $\mathbf{s^{-1}}$

A good first step in developing a neural network model is to look at the distribution of data. Model development also helps to analyse the data, a first model was developed which took 8 elements and the temperature strain rate, strain and a target of flow stress, to see how well the behaviour could be modelled before any of the inputs where altered.

These inputs contain all the information necessary to produce predictions for any combination of composition, temperature and strain rate, but it is important to realise that the complexity of the neural network has been limited by the regularisation which was designed to prevent over−fitting. The model is then a generalisation of the trends in the data, It was seen in making predictions that the experimental stress was not completely modelled by the predictions and lay outside the predicted error bars. The next models to be developed used inputs of 1/T, ln(strain), ln(strain rate) plotted against ln(stress).

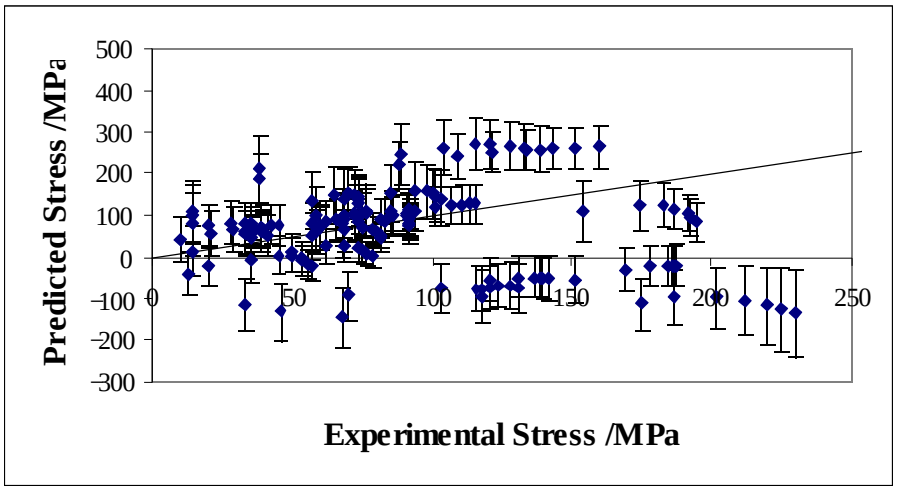


Figure 7.1, Model 1 Prediction of Steel 1 data

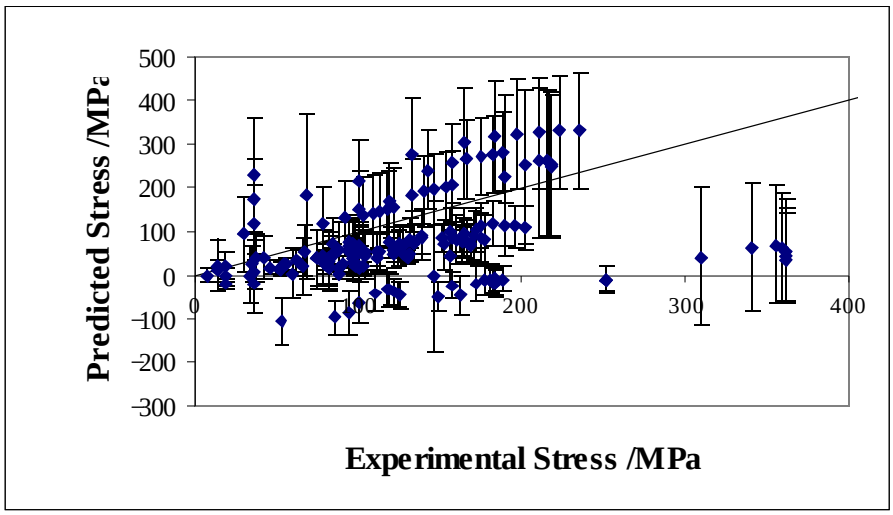


Figure 7.2, Model 1 Prediction of Steel 2 data

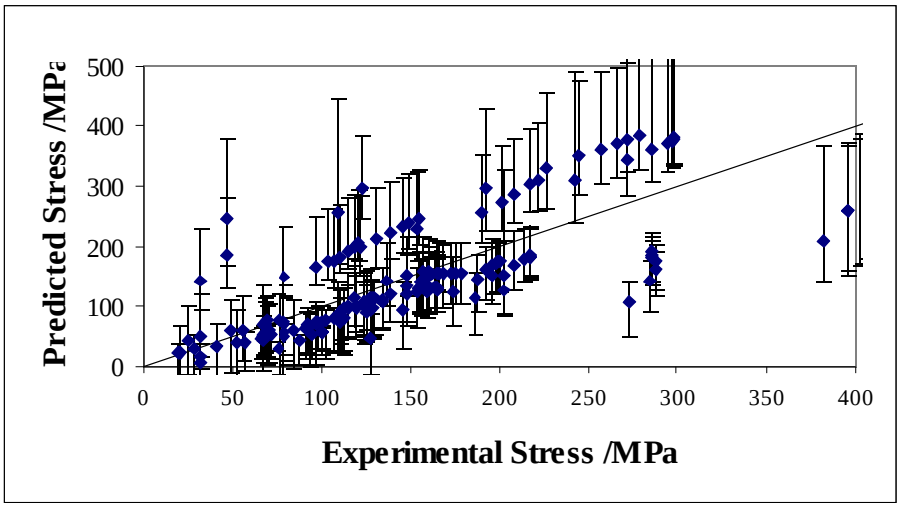


Figure 7.3, Model 1 Prediction of Steel 3 data

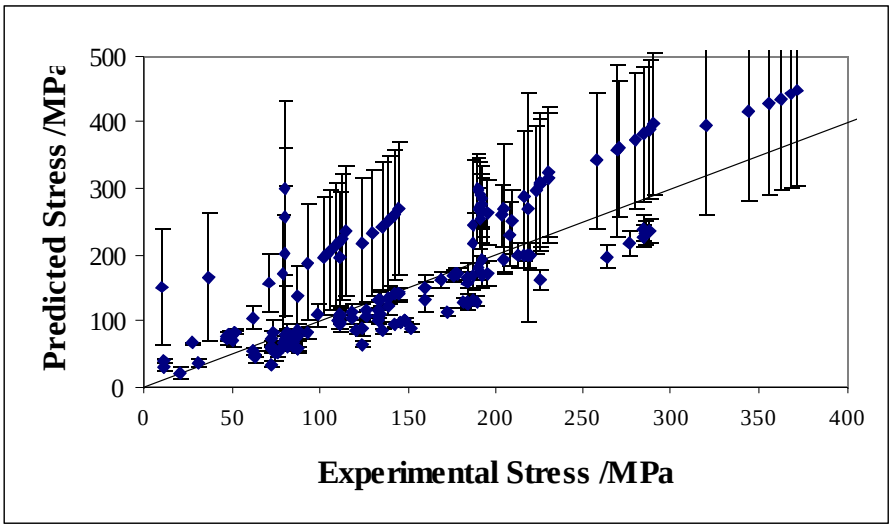


Figure 7.4, Model 1 Prediction of Steel 4 data

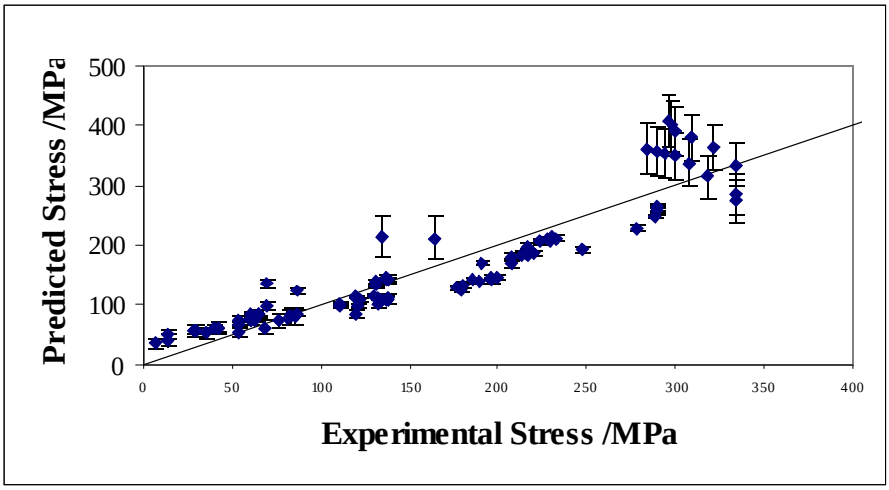


Figure 7.5, Model 1 Prediction of Steel 5 data

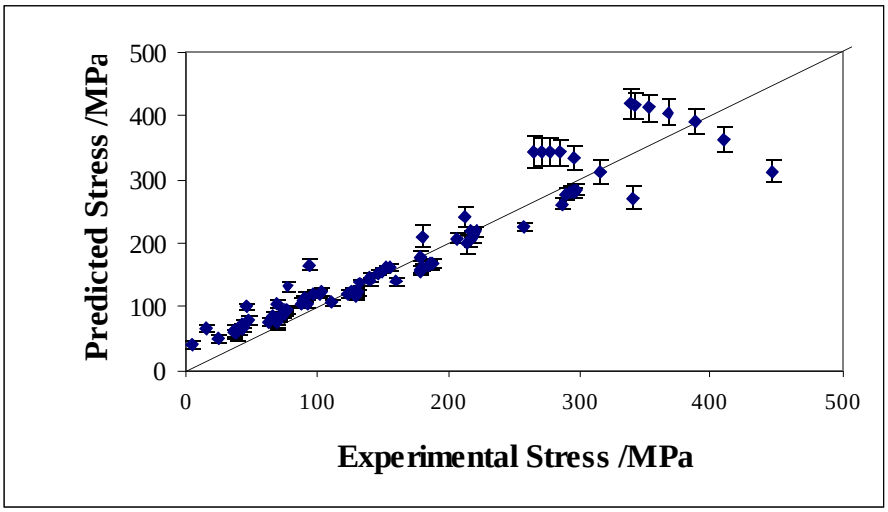


Figure 7.6, Model 1 Prediction of Steel 6 data

Figures 7.1−7.6 show the performance of model 1 in predicting the stress for data not included in the database. Steels 1−4 data are from compression tests at temperatures from

700 to 1200ºC and strain rates from 0.5 s^{-1} to 140 s^{-1}. Steel 5−6 data is from the same temperature range but at strain rates of 2 s^{-1} and 20 s^{-1} only. The results are encouraging in that some stresses are correctly predicted, especially those were the temperature and strain rate are within the range of the training data. The predictions that differ widely from the experimental values are usually accompanied by large error bars. However some of the predictions have negative values of stress, and this has to be recognised as a nonsense prediction in terms of a model of a compression test. Large variations from the experimental values were predicted for steels 1 & 2 at low temperatures, at these temperatures steels 1 & 2 contain some ferrite and so the regime is well away from the experience of the model, which only contains data for the austenitic region.

Changing the inputs to the inverse of the absolute temperature, and logarithm values of strain, strain rate and stress improved the model. This should help in several ways; firstly it scales the values better since the strain rate varied over several powers of 10. More importantly power laws and exponential functions physically relate these values, as shown by the dislocation theory and phenomological equations described in the introduction, and so plotting these functions should allow the neural network to use these relationships. Model 2 had inputs of 1/T, alloying elements in wt%, ln(strain), ln(strain rate) and a target of ln(stress). Data values for high stresses above 350 MPa were removed because their were relatively few points above this stress, this should improve the accuracy of predictions since the output values are scaled between +/1 0.5. Data for strains above 0.7 were removed also because only a few data points were available and so the predictions were being biased towards these values.

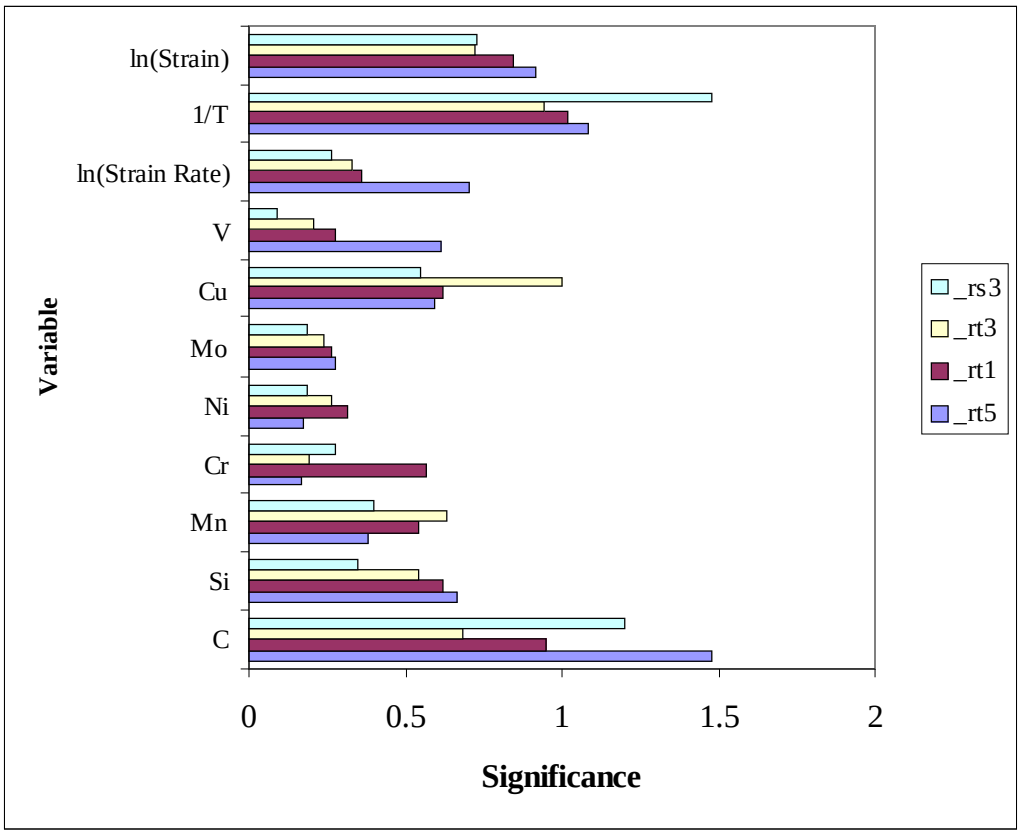


Figure 8, Bar−chart of the model perceived significance for each input for each sub−model in the committee, ?$_w$ values for model 2.

Next several models where developed using different inputs for composition. Weight percentages can be replaced with accurately calculated Atomic % to try and improve the ability to extrapolate to high alloy contents. It can be justified to use the square root of the atomic percentages since it has been reported that solid solution strengthening is related to the square root of the solute content. Theory suggests that the square root of the composition is proportional to the solid solution strengthening. Some success was obtained using inputs of alloy% and √alloy% for predicting the behaviour of the training set and the testing set, with the significance of √carbon being higher than that for carbon alone. This gave better predictions for alloys 1−6, with non of the stress predictions being negative as in model 1. Figure 9 shows that the models gave good predictions of the stress at higher strains, and stress levels than included in the training set.

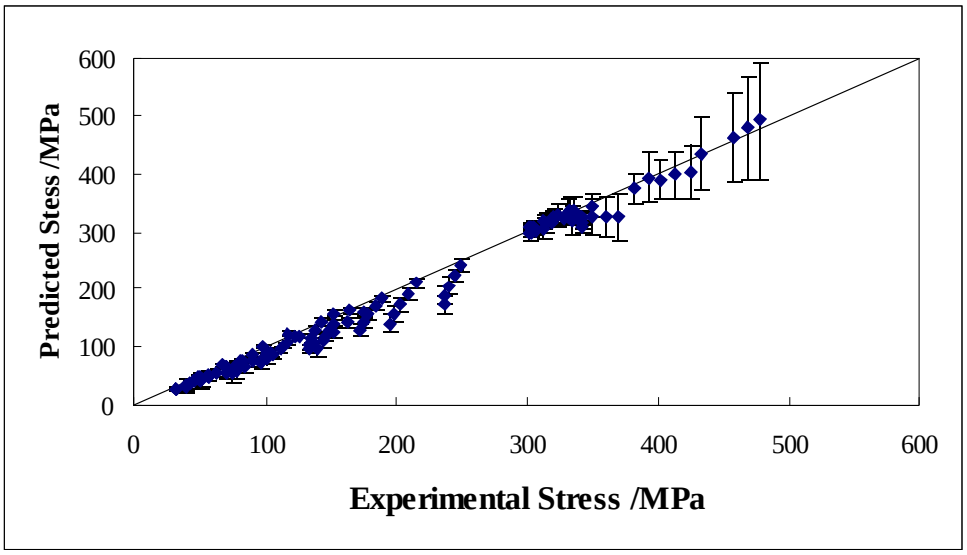


Figure 9a, model 3, predictions of high strain for compositions included in training set at Strains > 0.7

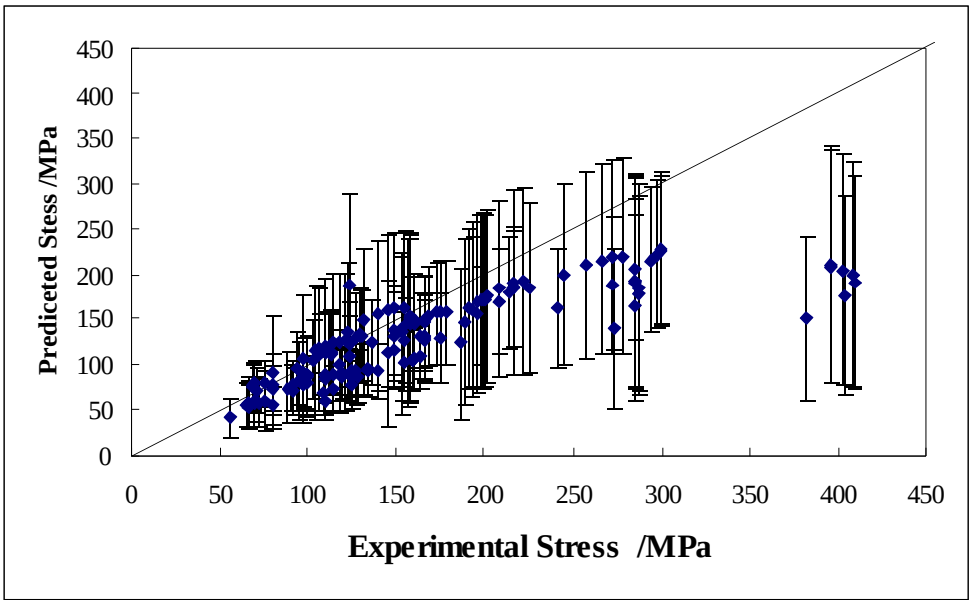


Figure 9b, model 3 predictions of Steel 3

Model 4 took the physically based variables as in model 2 and 3 of *1/T*, ln(strain rate), ln(strain) and ln(stress).

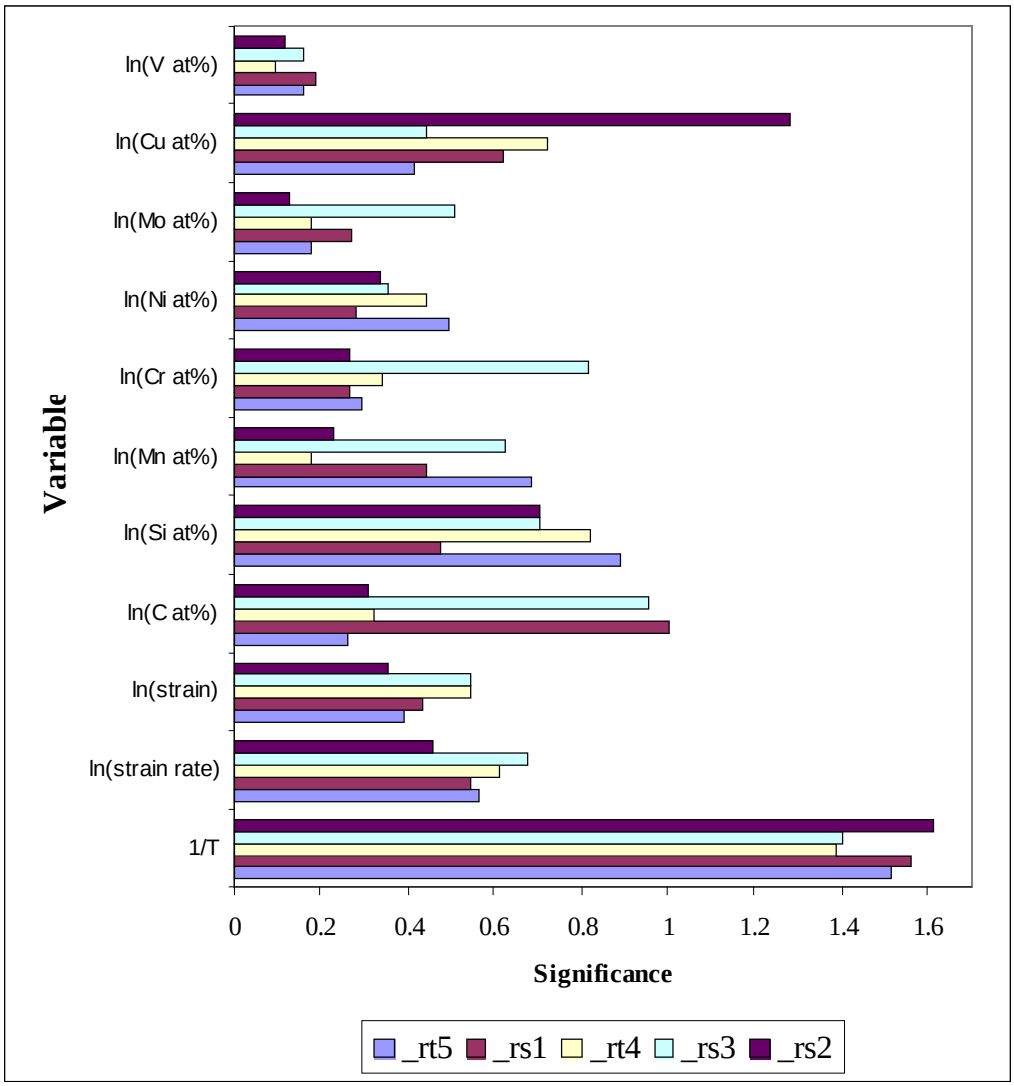


Figure 10, Significance of variables in model 4.

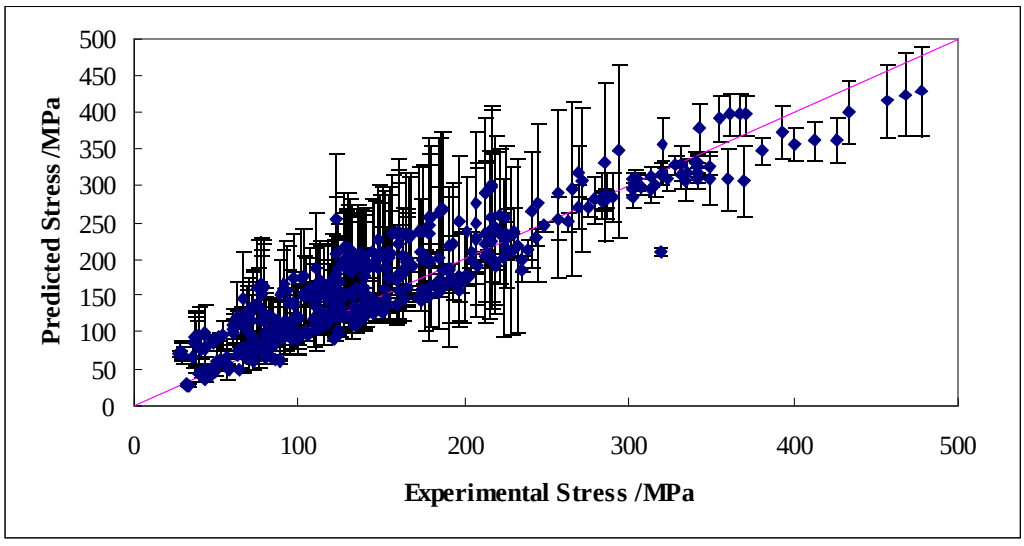


Figure 10, Steels 1−6 at high temperatures

Model 5 used the activation energy as an input, in the form of a *Q/RT* term in the model, calculated by the existing linear regression model proposed by Medina and Hernandez. Eqn. 7.4, [18].

Table 5, Calculated values of activation energy

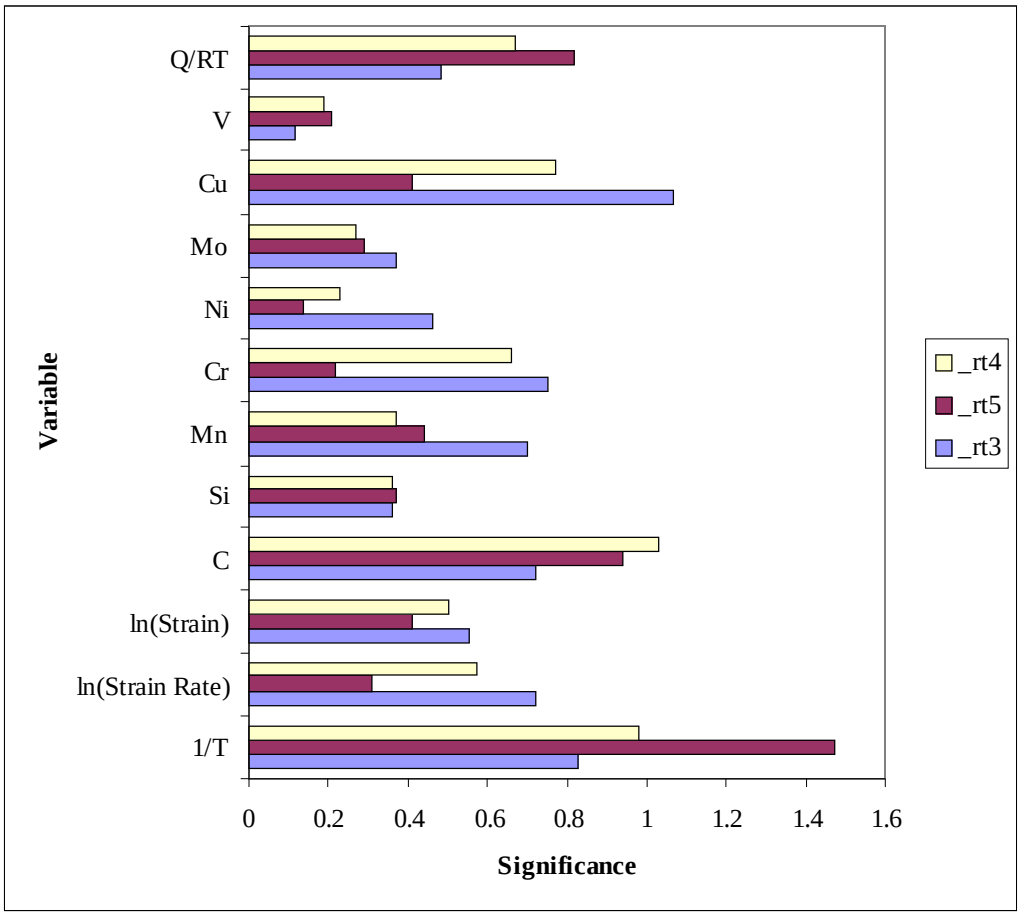


Figure 11a, submodels in the activation energy model

Vanadium was the only micro−alloying element that had enough data values to be significant to the models, this was disappointing because these have a effect through forming precipitates. The other alloying elements are assumed to be in solid solution in the model, they will have an effect by solid solution strengthening, and by changing the thermal activation energy.

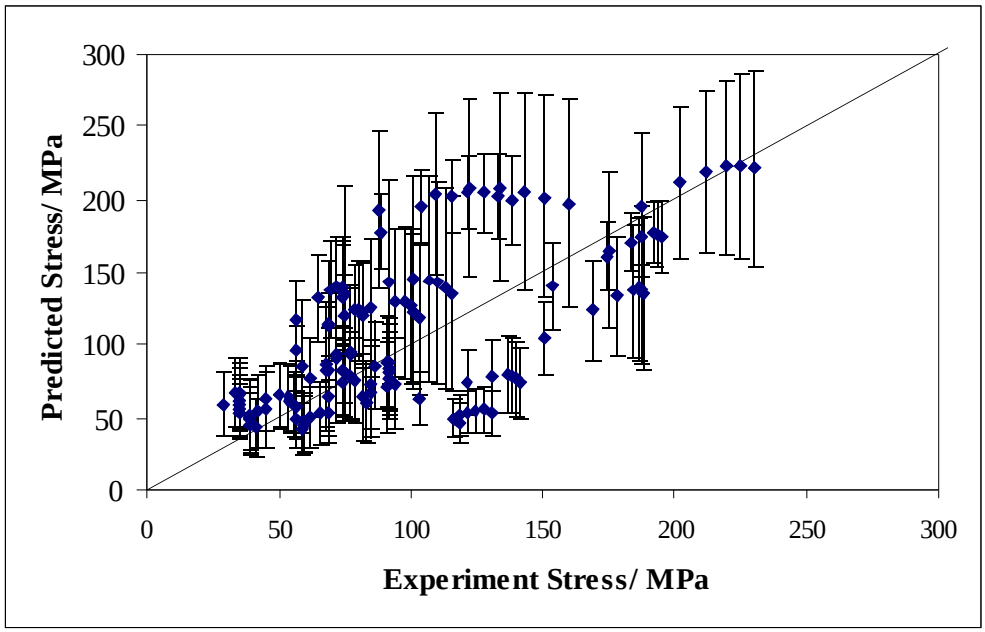


Figure 11b, Steel 1 Q model

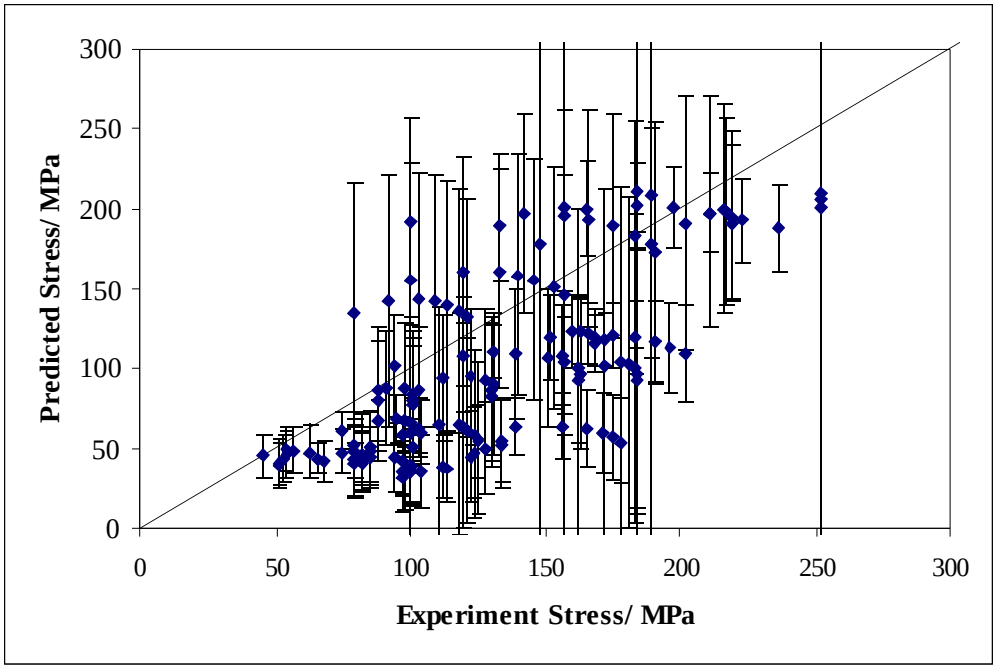


Figure 11c, Steel 2 Q model

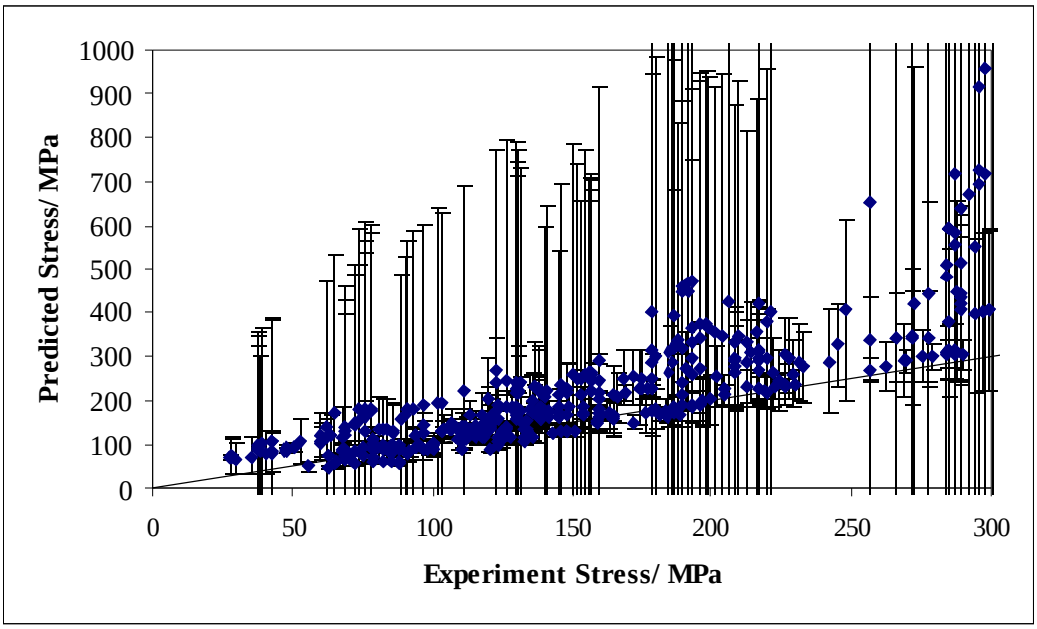


Figure 11d, Steels 3,4,5,6 Q model

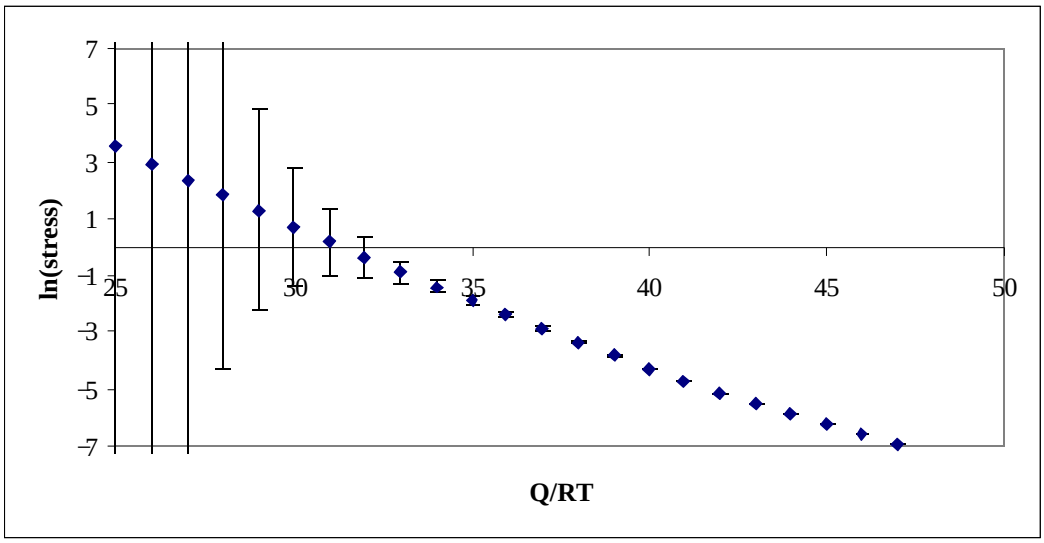


Figure 11e, Sensitivity of stress to Q, Grade78C, Temperature 1100ºC, strain rate 0.3s^{-1}, strain 0.1

Discussion

Several neural network models have been built from the data summarised in table 3 and table 5. All of the models were capable of modelling the behaviour demonstrated in the data set to a lesser or greater degree. The first model was used to demonstrate some of the trends that can be extracted from the model by changing one variable and keeping others constant. These trends may change with the position in the input space, this being a major advantage of the neural network models in comparison to the more traditional linear regression techniques.

Figure 4 shows the perceived significance of the inputs in the simple model which took the inputs to be the strain rate, temperature in ºC, carbon, silicon, manganese, copper, vanadium, chromium, molybdenum, chromium, and then nickel in order of significance. The trends identified by the model are demonstrated in figures 6.1−6.11b and are discussed in the experimental detail section (page 27). Conflicting reports of the effect of alloy additions have been reported, this could be due to differences in experimental technique, starting grain size, and also simply due to the differences in composition, temperature and strain rate.

Implicit in all models is assumption that strain can be expressed as a function of composition, strain rate, temperature and strain: that the strain can be expressed as a mechanical state function. An input that may be important in predicting the stress is the grain size, other models have incorporated the effect of change in grain size explicitly incorporated. Since the grain size was not known for the data used here, this may have caused an increased error in predictions of stress, especially at low strains. It would have been interesting to see the sensitivity of the flow stress to initial grain size, although in previous work Narayan *et al*.[25] concluded that the stress−strain curves were independent of "strain history".

Improvements were made to the simple model by attempting to make inputs physically more relevant. The equations for stress strain relationship, particularly equation 2.1, as well as the work hardening theory (which suggests stress is proportional to the square root of the strain), suggests a power relationship between the strain and the stress is appropriate. This means that if we make logarithmic values a straight line can relate be used to related the stress and strain. Also the temperature, or the inverse of the absolute temperature, and the strain rate both have a straight line relationship with logarithmic values of stress since equation 8.2 reduces to a power law relationship at low strains and an exponential relationship at high strains. Hot deformation takes place at high strains so the exponential form is appropriate.

Liu proposed an improved unification method *'since strain rate changes severely and causes too small a unified value'* since their strain rate changes by 10 times, they used a logarithm method, where:

$$\&= \frac{(4+\log\&)-0.95(4+\log\&_{min})}{1.05(4+\log\&_{max})-0.95(4+\log 4+\log\&_{min})}$$

In the method applied in this report the data was scaled using a logarithmic function that also had a physical basis. Figures 6.10b and 6.11b show logarithmic plots of predictions from the simple model. From the predictions the relationships could conceivable be

linear between ln(strain rate) and ln(stress) as well as ln(strain) and ln(stress). Although the neural network can cope with non−linear trends, changing to the physically more significant forms simplifies the regression task. The complexity of the model is limited by regularisation terms which penalise large weights and therefore the complexity is limited, the purpose is to prevent over−fitting, one consequence maybe that the modified models are able to generalise the trends more reliably than the simple model.

Teggart reported that the flow stress decreases with increasing temperature except in the two−phase ferrite plus austenite region were the reverse is true. Since the model was trained using data from the austenite region only, the model failed to predict the behaviour of testing steels 1 and 2 at low temperatures (700ºC), however large error bars correctly identified the poor predictions. The flow stress of other steel compositions at 700ºC could be more correctly predicted were the steel was austenitic at the temperature.

Further models were developed on the assumption that alloying additions would strengthen the austenite by solid solution strengthening, theory suggesting that the strengthening effect of each alloy is proportional to the square root of the percentage added. Model 3 was developed and included the alloys as atomic percentages and the square root of the atomic percentages, the model gave improved predictions for the test data seen in figure 9b, compared to figure 7.3. The values in the submodel showed a greater significance at ?at% carbon than at% carbon, this is intuitive if we look at figure 6.2 which shows the effect of alloy additions of carbon, while some of the other elements seemed to have little effect and the significance was not improved or in some cases reduced. The next model was altered so that the alloys were included as logarithmic functions, this resulted in better predictions, and made sense physically if the effects of alloy additions are by solid solution strengthening. It should also be better than including the inputs twice, as this may bias the data as well as possibly being unnecessary duplication.

The model were the alloy additions were included as logarithmic functions gave better predictions of the stress than the previous models as can be seen when comparing the predicted values to the experimental values in figure 10. The model behaved well behaved well in that it predicted no negative values of stress, which is an improvement on model 1. The error bars of the prediction usually encompassed the experimental values. The values do seem to be systematically higher than the experimental values for steels 1−6. This may be because the database of steels were more heavily alloyed, or explained by differences in experimental technique. When collating stress−strain curves from published results it is sometimes unclear how the results have been adjusted for example to compensate for heating effects during the deformation.

The fifth model included an empirical estimate of the activation energy, the error bars predicted for this model were much larger than for model 5, however figure 11e shows that predictions of the effect of changes in activation energy, with the correct physical relationship of Q/RT predicted by dislocation theory and the empirical models.

Conclusions

Neural networks are a general form of regression, and the results demonstrate how it is more appropriate than traditional linear regression techniques, because in physical systems the effect of changing one variable depends on the current values of the other

variables. For example in hot working the effect of changing carbon content depends upon the temperature and strain rate.

The quality of the neural network is highly dependent on the quality of the dataset, the results can be improved by careful selection of appropriate inputs.

The compositions available for training meant the neural network was limited in its applicability to the 8 elements of carbon, manganese, silicon, chromium, nickel, vanadium, copper, and molybdenum. Although it has proved the applicability of this kind of model a larger database would further improve the results and extend the range of validity, especially since micro−alloying elements such as niobium have an effect on flow stress response during hot working.

Predictions can be made from the models presented here, which can interpolate and extrapolate the stress−strain behaviour of austenite in hot rolling, the confidence in the prediction is indicated by the magnitude of the error bars.

References

[1] W. B. Morrison, J. A. Chapman, Philosophical Transactions of the Royal Society of London, (1976), A**2,** 1307, pp 289−303, *Controlled Rolling.*

[2] H. K. D. H. Bhadeshia, ISIJ International, 39, No. 10, pp 966−979, (1999), *Neural Networks in Materials Science.*

[3] S. B. Singh, H. K. D. H. Bhadeshia and D. J. C. Mackay, H. Carey and I. Martin, Ironmaking Steelmaking, (1998), **25**, pp 355.

[4] P. Korczak, H. Dyja and E. Labuda, Journal of Materials Processing and Technology, **80**−1, pp 481, (1998).

[5] C. Dumortier, P. Lehert, P. Krupa and A. Charlier, Materials Science forum, **284**, pp 393 (1998).

[6] P. Millytoski, Journal of Materials Processing Technology, (1998), **79**, pp 9.

[7] P. Millytoski, J. Larkiola and J. Nylander, Journal of Materials Processing Technology, (1996), **60**, pp 399.

[8] G. E. Dieter, Mechanical Metallurgy SI Metric Edition, pp 307−308, (1988) McGraw Hill, London.

[9] J. J. Jonas, C. M. Sellars, and W. J. McG. Teggart, *Metell. Rev.* , (1969) Rev. **14**, pp 1−24.

[10] J. Hollomon, Transactions AIME, (1947), **171**, p535, *The Mechanical Equation of State*.

[11] G. I. Taylor, (1934) Proc. Roy. Soc., **A145**, 362, *Mechanism of plastic deformation of crystals.*

[12] D. Hull and D. J. Bacon, Dislocations, pp 69, (1984), Pegamon Press Ltd, Oxford.

[13] C. Zener, J. H. Hollomon, Journal of Applied Physics, (1943), **15**, pp22−32, *Effect of Strain Rate Upon Plastic Flow of Steel.*

[14] W. J. McTegart and A. Gittins, The Hot Deformation of Austenite, J.B. Ballance Editor, pp 1−46, (1977), American Institute of Mining and Petroleum Engineers, *The hot deformation of Austenite.*

[15] Physical Metallurgy and the design of steels, F. B. Pickering, (1978) Applied Science Publishers, England.

[16] Steels, Microstructure and Properties, R. W. K. Honeycombe and H. K. D. H. Bhadeshia, (1995) Edward Arnold, London.

[17] M. J. Stewart, The Hot Deformation of Austenite, J.B. Ballance Editor, pp 47−67, (1977), American Institute of Mining and Petroleum Engineers, *Hot Deformation of C−Mn Steels from 1000 to 2200°F (600 to 1200°C) with constant true strain rates from 0.5 to 140 s^{-1}.*

[18] S. F. Medina and C. A. Hernandez, Acta Mater, 44, No. 1, pp 137−148, (1996). *General Expression of the Zener−Hollomon Parameter as a function of the chemical composition of low alloy and mircoalloyed steels.*

[19] C. M. Sellars, W. J. McG. Teagart, International Metallurgical reviews, (1972), **17**, p 1−24.

[20] K. Ghosh, Acta Met, (1980), **28,** p 1443−1465, *A Physically−based constitutive model for metal deformation.*

[21] C. G. Schmidt and A. K. Miller, Res. Mech. , (1981) **3**, p108−193.

[22]C. S. Hartley, R. Srinivasan. Transactions of ASME , **105**, p162−167, (1983), *Constitutive Equations for large plastic deformation of metals.*

[23] Universtiy of Cambridge Programme for Industry, (1995), *Modelling Phase Transformations in Steels, Bayesian Non, Linear Modelling with Neural Networks.*

[25] V. Narayan, R. Abad, B. Lopez, H. K. D. H. Bhadeshia and D.J.C. MacKay, ISIJ International, (1999), 39, No. 10, pp 999−1005, *Estimation of Hot Torsion Stress Strain Curves in Iron Alloys Using a Neural Network Analysis.*

[26] Y.−J. Hwu, Y.−T. Pan and J.G.Lenard, Steel Research, (1996), **67**, No. 2, pp 59−66, *A Comparative Study of artificial neural networks for the prediction of constitutive behaviour of HSLA and carbon steels.*

[27] L. X. Kong and P. D. Hodgson, ISIJ International, (1999), **39**, No. 10, pp 991−998, *The Application of Constitutive and Artificial Neural Network Models to Predict the Hot Strength of Steel.*

[28] J. Liu, H. Chang, T. Y. Hsu, X. Ruan, Journal of Materials Processing Technology, (2000), **103**, pp 200−205, *Prediction of the flow stress of high−speed steel during hot deformation using a BP artificial neural network.*

Not Specifically referenced

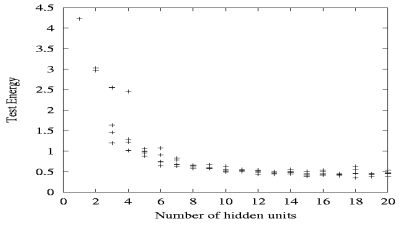
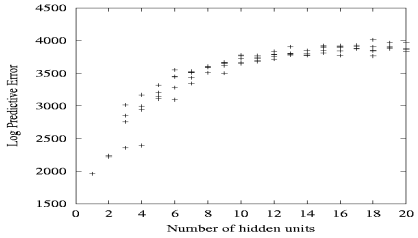
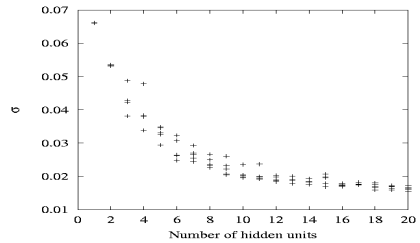
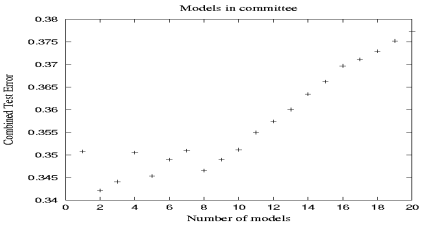
D. C. Collinson, P. D. Hogdson and B. A. Parker, pp 283, (1993), Modelling of Metal Rolling Processes, Institute of Materials, London.

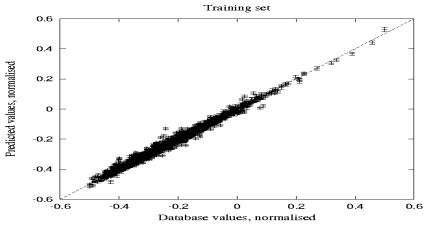
Applying Neural Networks, A Practical Guide, K. Swingler, (1996), Academic Press, London.

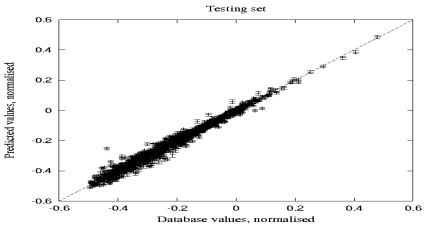
R. Colás, Journal of Materials Processing Technology, **62**, pp180−184, (1996), *A model for the hot deformation of low carbon steel.*

Alloy	**Calculated *Q*** **kJ**	***Alloy***	***Calculated Q*** ***kJ***

78C_railsteel	268	316	261
50Concast	268	20C	267
55Concast	268	16C_V	267
43Concast	268	95C	267
100C	267	25C	267
55C	267	12C	270
15C_A	276	Leaded steel	330
En16	266	08C	267
En31	267	45C	267
En52	267	Steel1	267
VNiCu	282	Steel2	267
En45	267	Steel3	267
En40	266	Steel4	267
En25	266	Steel5	267
304	267	Steel6	267

a) TE	b) LPE
	
c) Sigma	d) Committee
	
e) Training set − submodel rt5	f) Testing set − sub−model rt5





Variable	Range	Mean	Standard Deviation
C (wt%)	0.07−1.06	0.33	0.28
Si (wt%)	0−3.74	0.44	0.78
Mn (wt%)	0.28−1.58	0.86	0.43
Cr (wt%)	0−18.6	1.97	4.86
Ni (wt%)	0−12.04	0.84	2.4
Mo (wt%)	0−2.26	0.12	0.36
Cu (wt%)	0−0.32	0.03	0.08
V (wt%)	0−0.15	0.02	0.04
Strain rate (s^{-1})	0.1−100	21.03	32.47
Temperature (ºC)	800−1300	27.18	132.05
Strain	0−1.75	0.32	0.26
Stress MPa	6−478	131.08	65.75